ATLANTIS
Journal of the Spanish Association of Anglo-American Studies
44.2 (December 2022): 45-64
e-ISSN 1989-6840
DOI: http://doi.org/10.28914/Atlantis-2022-44.2.03

Factors Affecting Pronunciation Accuracy in English as a Foreign Language: The Case of Spanish-Catalan Intermediate Learners

Lucrecia Rallo Fabra
Universitat de les Illes Balears
lucrecia.rallo@uib.es

In the present study, we investigate the factors that influence pronunciation accuracy by Spanish-Catalan bilinguals learning English as a foreign language (EFL). A group of intermediate EFL learners were recorded producing a series of cognate and non-cognate words in two different conditions: a reading aloud task and a delayed repetition task. In the reading aloud task, the target words were presented as visual prompts in a carrier phrase. In the delayed repetition, words were presented aurally and participants repeated them after a two-second delay followed by an audio prompt. The words were phonemically transcribed using PHON. As expected, task condition and cognate-status influenced the percentage of correct vowels and consonants. The number of errors was significantly higher in the reading aloud condition than in the delayed repetition condition. Cognates exhibited more pronunciation errors than non-cognates. In contrast, learners' vocabulary-size and lexical frequency did not have a clear impact on the results. These findings suggest that focusing on spelling might interfere with the way EFL learners process the phonological forms of words.

Keywords: pronunciation; orthography; cognates; vocabulary-size; lexical frequency; EFL

. . .

Factores que influyen en la pronunciación del inglés como lengua extranjera. El caso de los aprendices de L1 español-catalán a nivel intermedio

En este estudio se investigan los factores que contribuyen a la pronunciación del inglés como lengua extranjera por parte de hablantes bilingües de castellano-catalán como L1.

Un grupo de aprendices de inglés de nivel intermedio fueron grabados mientras producían cognados y no cognados en dos condiciones experimentales: una tarea de lectura en voz alta y una de repetición retardada. En el primer caso, las palabras se presentaron de forma visual en el contexto de una frase portadora. En la repetición retardada los participantes escuchaban las palabras a través de auriculares y las repetían después de la señal sonora que se reproducía automáticamente después de dos segundos. Dos jueces expertos transcribieron fonéticamente las palabras mediante la herramienta PHON. Como era de esperar, tanto la tarea como la condición de cognado influyeron en el porcentaje de vocales y consonantes producidas correctamente por los participantes. El número de errores fue significativamente mayor en la tarea de lectura en voz alta en comparación con la repetición retardada. Los cognados presentaban más errores de pronunciación que los no cognados. Sin embargo, ni el tamaño del vocabulario ni la frecuencia léxica tuvieron un impacto claro en los resultados. Concluimos que es muy probable que la ortografía interfiera en el procesamiento fonológico de las palabras.

Palabras clave: pronunciación; ortografía; cognados; tamaño del vocabulario; frecuencia léxica; inglés lengua extranjera

1. Introduction

Pronunciation teaching does not receive much attention in the English as a foreign language (EFL) curricula in Spain. This is surprising if we consider that pronunciation is fundamental for successful communication (Levis and Grant 2011, among others) and students often state that the time teachers devote to pronunciation activities is not sufficient (Foote et al. 2016; Fouz-González 2020). A survey about pronunciation teaching at different education levels in ten European countries (Henderson et al. 2015) revealed the limitations of many teachers to teach this skill. When they were undergraduate/postgraduate students, most of Henderson's respondents took theory-based English phonetics and phonology courses combined with practical training to improve their own English pronunciation. However, most of them acknowledged that they did not receive specific training on how to teach pronunciation at primary, secondary or tertiary levels. This limitation had a clear effect when the weight given to this skill in classes taught by the respondents was analyzed. On average, the teachers surveyed did not spend more than 25% of the time on pronunciation practice. Some respondents gave priority to successful communication, even if this went hand-in-hand with pronunciation errors. In the case of state-run schools in Spain, students receive massive written input but limited auditory input. When students have a chance to practice speaking through oral presentations and dialogues, corrective feedback is not always provided by the teacher, who, like the students, is also a non-native speaker. This results in pronunciation errors that fossilize over the learner's educational career and seriously compromise intelligibility (Gallardo del Puerto et al. 2009; Rallo-Fabra and Juan-Garau 2011; Rallo-Fabra 2015).

According to Levis (2005), the *intelligibility principle* has dominated research and pedagogy in second language (L2) pronunciation in recent decades. This approach advocates that communication is what matters and attempting to achieve native-like pronunciation is an "unrealistic burden for both teacher and learner" (Levis 2005, 370). As shown in the strong empirical evidence provided by Derwing and Munro (2015), L2 speech can be intelligible to the listener and yet be strongly accented. Studies exploring pronunciation by Spanish-Basque young learners in classroom settings provide additional evidence. For instance, Gallardo del Puerto et al. (2009) investigated language gains in two educational contexts, i.e., conventional formal instruction (FI) versus content-and-language integrated learning (CLIL). Pronunciation was assessed along three dimensions: a) foreign accent (FA); b) FA intelligibility; and c) FA irritation in a narrative task. As expected, expert listeners judged the CLIL students' pronunciation as more intelligible and less irritating than the FI students' pronunciation. However, no significant differences in FA ratings were found between the two groups of learners, suggesting that EFL speech at intermediate levels was intelligible even when strongly accented. Similar results were found by Rallo-Fabra and Juan-Garau (2011). Spanish-Catalan CLIL students were perceived by expert judges as easier to understand than their FI peers, although the two groups did not significantly differ in terms of the

perceived FA. Gallardo del Puerto et al. (2009) concluded that many pronunciation errors could have been an artifact of the task since only read-aloud speech samples were used, and therefore many pronunciation errors could have been "spelling-induced," i.e., caused by the mismatch between English graphemes and phonemes.

In the following section we review the findings of previous theoretical and empirical studies that have explored the role of different factors in L2/FL pronunciation. We then provide empirical data of pronunciation outcomes by a group of intermediate Spanish-Catalan learners in an instructional setting and test the relative weight of various factors, namely, orthography, cognate-status, lexical frequency and vocabulary-size.

2. Factors Influencing L2 Pronunciation in Instructed Settings

A review article by Thorsten Piske et al. (2001) explored the factors that contribute to FA in L2 speech production. Although the review was based on L2 learners in a naturalistic setting, some of the factors reported by the authors are also relevant in instructed settings, such as amount and quality of formal instruction, motivation and language-learning aptitude. More recently, pronunciation learning by L2 learners has been found to be subject to a high degree of individual variability. Even when learners share the same native language (L1) background and proficiency level, they do not always share the same difficulties (Derwing and Munro 2015). Among the factors reported in the literature known to affect the pronunciation of individual sounds (segments) are orthography, lexical frequency and phonological context.

2.1. Orthography Effects

On the basis of orthography, languages can be divided into transparent and opaque (Erdener and Burnham 2005; Erdener 2016). Transparent languages such as Spanish have a one-to-one correspondence between sounds and letters. In contrast, languages such as English have an opaque orthography and, as such, a given spelling might have more than one pronunciation. For instance, the <*th*> grapheme is pronounced /θ/ in *thought* but /ð/ in *though*. Similarly for vowels, the grapheme <*i*> is pronounced /aɪ/ in *high* but /ɪ/ in *hit* and <*a*> is pronounced /æ/ in *cat* but /eɪ/ in *name*. Orthographic complexity has been found to vary across different writing systems (Chang et al. 2016) and it can be operationalized in *z* scores.[1] For instance, a language like Chinese, which ranks the highest in the list of orthographic complexity, has a *z* score of 3.79. English and Spanish have *z* scores of -0.50 and -0.61, respectively, indicating that English orthography is more complex than Spanish.

Orthographic input is the main source of L2 input in instructed settings (Hayes-Harb and Barrios 2021). It follows that, if learners are literate when they deal with L2

[1] *z*-scores are a data normalization procedure to center the values around 0. To calculate them, "we take each score and substract from the mean of all. Then we divide the resulting score by the standard deviation (SD) to ensure that data have a SD of 1." (Field 2012, 26).

written input, the phonological and orthographic representations of words in the L1 and L2 will necessarily interact with each other. Specifically, "mismatches between the grapheme-phoneme correspondences between the native and second languages (referred to here as incongruency) can interfere with learners' production of L2 words" (Hayes-Harb and Barrios 2021, 305). In the same line, the work of Bassetti and Atkinson (2015) exemplifies these interactions, and it is relevant to our study because, just like Spanish and Catalan, Italian is a transparent language. Furthermore, the experimental conditions of Bassetti and Atkinson's (2015) study are similar to ours, i.e., intermediate learners of English and elicitation methods that consisted in reading aloud and repetition tasks. The Italian learners produced pronunciation errors that were mostly orthography-induced, such as the production of silent letters in *debt* *[debt] or the lengthening of vowels in words that had double letters like *door* among others.

Just like their Italian peers, Spanish-Catalan children learning EFL in classroom settings are often exposed to considerable written input before they are able to speak the language. This has serious consequences for their pronunciation skills. When reading English, students tend to apply the rules of their L1 (one-letter-one-sound correspondence), which triggers orthography-induced pronunciation errors understood as "cases in which a speaker aims to produce an utterance, but as a result of a lack of full control over its segmental or suprasegmental structure, produces something else instead" (Derwing and Munro 2015, 57). In the absence of corrective feedback, these pronunciation errors are often internalized by the learner and are most likely to fossilize throughout the lifespan.

2.2. Word Familiarity, Lexical Frequency and Cognate Effects

Recent studies examining pronunciation learning in instructed settings have found that word familiarity may have an impact on the pronunciation accuracy of L1-English L2-French learners (Woore 2018). If learners are faced with a familiar word, they might retrieve pronunciations stored in long-term memory. However, if they encounter an unfamiliar word, they will probably need to match the grapheme(s) with the corresponding L2 phoneme applying the rules of the L1. It follows that if these do not match the L2, accented pronunciation will result. Similarly, Trofimovich et al. (2012, 183) provide evidence that lexical frequency counts provided by large corpora are reliable estimates of a learner's encounters with a language. The authors argue that learning accuracy in L2 phonology can be predicted from the lexical frequency of the input lexical items they receive—the more often the learner hears a word, the more accurately it will be produced. Another study by Vokic (2011) exploring the production of the English flap by Spanish speakers in an immersion setting showed that high-frequency words triggered less errors than low-frequency words.

Cognates are words in different languages that have originated from a common parent word (Schmitt and McCarthy 1997). Cognate vocabulary can accelerate the process of vocabulary learning if the L1 and the L2 share lexical similarities. However, the typological difference between the two languages will determine whether learners

will benefit from this cross-linguistic similarity of cognate words. If the languages are typologically close, such as Spanish and Catalan, learners will accept cognates as equivalent in form and function, but if the two languages are perceived as distant, they might not be aware of certain formal and functional similarities between them (Schmitt and McCarthy 1997; Ringbom 2007; Otwinowska-Kasztelanic 2009). For instance, a Spanish learner of Catalan will probably guess that Catalan *curt* is the equivalent of Spanish *corto* in form and meaning, but a Spanish-Catalan learner of English might not be aware that *gains* is the English equivalent of Spanish *ganancias* or Catalan *guanys*.

In the case of L2 phonology, the effects of cognate status on segmental speech production have been extensively investigated (Amengual 2012, 2016; Mora and Nadeu 2012; Rallo-Fabra 2015; Gorba et al. 2021). Two of these studies involve subjects who are bilingual in Catalan and Spanish, two languages of Romance origin that are typologically close. For instance, Mora and Nadeu (2012) found that Catalan-Spanish bilinguals produced less native-like instances of /ɛ/ in Catalan words that had a cognate in Spanish such as *beca* ("grant") or *guerra* ("war"). In contrast, non-cognates such as *pèsol* ("pea") or *res* ("nothing") exhibited more target-like instances of /ɛ/ (measured acoustically). More recently, Amengual (2016) reported similar findings for the mid-vowel /ɔ/, which again was found to be higher in the acoustic space in cognates relative to non-cognates, thus triggering a merge between /o/ and /ɔ/. When performing a lexical-decision task, incongruent cognates such as *sol* ("sun"), which have a different vowel in Spanish /sol/ and Catalan /sɔl/, increased cross-linguistic interference between the two vowels, causing a higher percentage of errors in the lexical recognition process.

Similar conclusions were reached in studies involving Spanish-English bilinguals and Catalan learners of English (Amengual 2012; Rallo-Fabra 2015). Spanish-English bilinguals and Spanish EFL learners produced English /t/ and /p/ with more Spanish-like VOT values in cognates than in non-cognates. Similarly, late Spanish-English bilinguals living in the U.S. produced more reduced vowels in unstressed syllables if the target word was a non-cognate (Rallo-Fabra 2015).[2] In other words, cognate words such as *permit*, triggered less reduced vowels and more accented production. Taken together, the results of these studies suggest that when cognates share the same spelling in the L1 and the L2 but have incongruent pronunciations, e.g. *bilingual* in English /baɪˈlɪŋgwəl/ but *bilingüe* /biˈliŋgwe/ in Spanish, the tendency is to apply the letter-to-sound rules of the L1, resulting in accented pronunciation in the L2.

2.3. Vocabulary Size

Many studies have investigated the relationship between vocabulary-size and listening and/or comprehension ability. In contrast, studies relating vocabulary-size and speaking

[2] *Late bilinguals* is used refer to people exposed to the L2 when the L1 has already been acquired.

skills are scarce. Vocabulary competence can be tested using receptive- and productive-vocabulary tests. The former measures "the set of words that a learner recognizes and understands," while the latter refers to "the set of words that a learner uses when writing or speaking" (Mairano and Santiago 2020, 145). The authors examined the relationship between receptive and productive vocabulary on the pronunciation of L2-French by L1-Italian learners. Pronunciation was measured impressionistically by means of native judgements of FA, as well as with acoustic distance between vowel contrasts and fluency measures. They found that measurements of productive vocabulary were positively correlated with fluency, but no correlation was found between receptive vocabulary and any of the objective or subjective pronunciation measures.

In the perception domain, Best and Tyler (2007) hypothesized that a large L2 vocabulary may push learners "to attune to articulatory, phonetic and phonological differences in the L2 that have been previously ignored in the L1" (Best and Tyler 2007, 53). In other words, the more experienced learners become, the more sensitive they are to sounds that are not contrastive in the L1 but are crucial to differentiate pairs of confusing words in the L2 such as *hit* and *heat* or *cup* and *cop*. In this same vein, Bungaard-Nielsen et al. (2011a, 2011b) tested the hypothesis that vocabulary size might be driving the process of *rephonologization*—understood as the adaptation of L1 phonological categories to accommodate L2 categories, either by adding new ones or modifying existing ones to serve both languages. Following Nation (2006), they divided a group of Japanese learners of English into two groups, Japanese High-Vocabulary (HV) and Low-Vocabulary (LV). The threshold was set at 6,000 words, since this is the minimum number of word families needed to understand spoken English effectively. They found that vocabulary-size influenced both the perception and the production of L2 sounds, suggesting that lexical expansion goes hand-in-hand with phonological development.

2.4. Task Effects

As noted by Munro (2008, 202), one of the issues faced by researchers in L2 pronunciation is choosing a suitable elicitation method. He argues that reading aloud material may trigger mispronunciations caused by word familiarity or orthographic interference. He therefore recommends complementing this task with a "delayed repetition task," which reduces "the short-term recall of the model voice that might otherwise allow close imitation" (2008, 202).

Elicited imitation tasks (EIT) are a common method used to measure oral proficiency, as reported by Gaillard and Tremblay (2016). They argue that EIT relies on working memory and it can be adapted to test populations with varying levels of achievement by means of lengthening or shortening the number of words to be repeated. They reported correlations between EIT performance and cloze test scores as well as learners' experience with L2-French, thus providing sufficient evidence of the validity and reliability of EIT for L2 proficiency assessment.

3. The present study

3.1. Aims and Research Questions

In light of the findings reviewed above we set out to examine the relative weight of four factors—task, cognate status, vocabulary-size and lexical frequency—that may influence the pronunciation accuracy of intermediate Spanish-Catalan EFL learners in a classroom setting. Specifically, we address the following research questions:

RQ1: Does elicitation condition (reading aloud vs. delayed repetition) influence Spanish-Catalan EFL learners' pronunciation accuracy of English words? We predict that participants will be more accurate in the delayed repetition (DR) condition because in the reading aloud (RA) condition, orthography will interfere with the phonological representations of the target words and thus trigger more pronunciation errors.

RQ2: Will cognates be bound to more pronunciation errors than non-cognates? The literature reviewed above suggests that cognates are more prone to pronunciation errors than non-cognates because learners tend to transfer the phonemic patterns of the L1 words onto the L2 words.

RQ3: To what extent will vocabulary-size influence pronunciation accuracy at the individual level? On the basis of the findings of Bungaard-Nielsen et al. (2011a, 2011b) and Mairano and Santiago (2020), learners with larger vocabulary sizes should be more accurate than learners with smaller vocabulary-sizes in producing the target words.

RQ4: Can FL segment production be predicted by lexical frequency and/or word familiarity? We predict that high frequency words will be less challenging for learners because the more they hear a word, the more target-like its production will be (Vokic 2011; Trofimovich et al. 2012).

3.2. Participants

Twenty-three Spanish-Catalan, intermediate-level EFL learners with an average age of thirteen years participated in the study. The participants were Spanish-Catalan bilinguals in the third grade at two state-run secondary schools in Majorca (Balearic Islands). All the participants completed a vocabulary-size level test (Nation 2010) to determine their proficiency levels in English. A subset of the participants (n = 10) also completed a specific vocabulary test in which they were asked about the meaning of the target words and whether they knew how to pronounce them.[3]

Only four participants out of twenty-three scored close to or higher than 6,000 words, suggesting that the majority of the participants had vocabulary sizes that would make spoken English comprehension challenging. Following Bundgaard-Nielsen et al. (2011a,

[3] Due to internal organization issues at one of the participating schools, responses to the specific vocabulary questionnaires could not be obtained for all participants.

2011b), we divided the participants into high (HV) and low vocabulary (LV) groups using a median split (HV: n = 11, $M_{vocabulary}$ = 4618; LV: n = 12, $M_{Vocabulary}$ = 1841).

3.3. Speech Materials and Elicitation Procedure

A total of 20 cognates and 20 non-cognates were selected from the vocabulary bank of a project about civil rights in which the students had participated (see table 1). Four extra words were added to familiarize the participants with the task but were not used in the analysis. Word frequency for each of the target words was computed with the SUBTL_CD index from *Lexique* (Gimenes and New 2016), which renders a normalized word frequency index as a percentage. This corpus was preferred over other more commonly used corpora such as the British National Corpus (BNC) or the Corpus of Contemporary American English (COCA) because its frequency counts are based on *Twitter*, blog posts and newspapers, which was judged to be the kind of input learners had access to on social networks. In addition, the fact that the SUBTL_CD index deals with percentages facilitates correlating this variable with the DR and RA ratios, which range from 0 to 1. In contrast, BNC and COCA word frequencies are reported in number of words per million, which makes correlations methodologically problematic because we would have been comparing measures that are quantitatively different.

The participants were recorded individually in a quiet room at the school premises in two separate sessions. There was a two-month interval between the two sessions and no specific pronunciation training was provided prior to the recordings. Based on Munro's (2008) recommendation of not basing pronunciation assessment on read aloud materials only, two elicitation methods were used for the data collection, a reading aloud (RA) task and a delayed repetition (DR) task. In the first session, the participants read the target words, which were presented first in the carrier phrase *I say* ___ *this time*, and then in citation form. Only the former were considered in the analysis because many participants self-corrected their production when the word appeared in isolation. In the second session, the participants heard the target words produced by a native British English speaker over headphones, and they were instructed to model them after they heard the sound of a bell. The orthographic form of the word was not shown. This method had been previously used in L2 speech production studies (Piske et al. 2001, among others).

Table 1. Target words produced by Spanish-Catalan EFL learners. The SUBTL_CD indexes indicate word frequency ranging from 0% to 100%

Cognates	SUBTL_CD	Non-cognates	SUBTL_CD
humanitarian	0.89	homeless	3.92
racism	0.62	country	39.32
bilingual	0.17	challenge	8.35
charity	6.41	people	94.05
conflict	4.02	journalist	2.21

Cognates	SUBTL_CD	Non-cognates	SUBTL_CD
concentration	2.59	village	9.23
refugee	0.87	peace	22.26
prisoner	7.52	money	67.32
doctor	44.15	disease	9.78
soldier	9.9	border	5.51
medicine	11.92	war	30.35
government	18.16	speech	12.99
organization	0.88	Jewish	4.17
hypothermia	0.39	weather	12.29
hospital	31.13	survivor	1.82
malnutrition	0.27	environment	5.5
politician	1.65	world	73.84
negotiate	0.05	rights	99.19
diary	2.8	funds	3.06
international	5.83	hunger	2.83

3.4. Pronunciation Accuracy

A total of 920 words were phonemically transcribed from audio files using the *PH0N* software (Rose and MacWhitney 2015). The transcribers were a native speaker of British English living in Spain and a Spanish-Catalan speaker with a C2 level of English. Both raters were university lecturers teaching EFL courses and both had expertise in the phonemic transcription of L2 speech. Due to the large quantity of the speech material, each rater transcribed 50% of it. To ensure that there were no discrepancies in the criteria applied by the two transcribers, each one of them carried out a blind transcription of a randomized list of words such that 20% of the total sample was transcribed by both raters. Following Field (2012), intraclass correlation analyses were conducted to ensure that there was agreement between the two transcribers. The absolute number of correct consonants and vowels of 175 target words blindly transcribed by the two raters were correlated separately. The resulting Pearson correlations obtained were $r = 0.685$, $p < 0.001$ for the vowels and $r = 0.948$, $p < 0.001$ for the consonants, indicating that both raters agreed on the number of vowels and consonants correctly transcribed.

4. Results

Given the high variability of the target words in terms of number of syllables and phonetic contexts, different types of errors were observed. We found many instances

of consonant cluster simplification such as */ˈʤɜːnəlɪs/ for *journalist* or */ˈrasis/ for *racism*. Quite often learners added an epenthetic vowel to facilitate the pronunciation of complex consonant clusters, as in *government* */ɡoˈvernament/ and *speech* */esˈpitʃ/. Deletion of /h/ in initial position was also common, such as in *hypothermia* */ipoˈtermja/ and *hospital* */osˈpɪtal/. The influence of L1 phonology was probably responsible for spirantization of /d/ in middle position, for example in *medicine* */ˈmɛðisin/ and *border* */ˈborðər/.

As for vowels, the learners exhibited difficulties with production of segments that did not have a similar counterpart in the L1, such as the lax vowels /ɪ/ and /ʊ/. The absence of vowel reduction in unstressed syllables in words such as *people* */ˈpipol/ and *malnutrition* */ˌmalnuˈtrɪʃon/ was also a common trend. It is important to note that many pronunciation errors in the RA were spelling-induced mispronunciations like in *country* */ˈkoʊntri/ or *money* /ˈmoneɪ/. Some of the errors in the DR condition involved renditions that are quite distant from the native form, the most striking examples being */ˈʤeneɾist/ for *journalist*, */proˈteɪʃjan/ for *politician*, */ˈkɪwɪʃ/ for *Jewish* and */faʊnd/ for *funds*. Interestingly, the cognate effects surpassed the segmental level, some of the errors involved wrong stress placement, as is the case of *refugee* produced as */rɛˈfjuʤi/ matching the stress pattern of *refugi* /rɛˈfuʤi/ ("shelter") in Catalan.

To analyze the pronunciation errors quantitatively we ran a phonological error analysis of the transcribed words using *PHON* software, which yields percentages of correct segmental production separately for vowels and consonants. Pronunciation accuracy of the target words was thus operationalized by means of two dependent variables, the percentage of aligned correct consonants (%APCC) and the percentage of aligned correct vowels (%APVC). These are normalized measures of phoneme production calculated from the proportion of consonants/vowels in a given word and the number of consonants/vowels correctly produced in each word by a given speaker. These measurements have been used previously in developmental studies of L1 phonological acquisition in typically-developing children and in children with speech disorders (McAllister-Byun and Rose 2016). Due to the bimodal distribution of the results, we ran separate *Mann-Whitney U* tests to test the effects of the independent variables *task condition* and *cognate-status*. Such analyses are recommended for categorical data and/or when the data are not normally distributed (Field 2012).

4.1. Effect of Task Condition

The mean percentages of APVC and APCC tallied across word type and task condition are shown in table 2. Overall, task condition affected both the production of vowels (%APVC) [$U = 353.72$, $z = 8.08$, $p < 0.001$] and consonants %APCC [$U = 421.23$, $z = -2.38$, $p = 0.017$]. In both cases, segment production was more accurate in the DR task than in the RA task.

4.2. Effect of Cognate Status

In the DR task, %APCC of cognate words did not differ significantly from that of non-cognates [U = 107.73, z = 1.9, p = 0.054]. However, the %APVC in cognate and non-cognate words did differ significantly [U = 117,77, z = 4.5, p < 0.001], indicating that the participants produced a significantly higher number of correct vowels in the non-cognates than in the cognates.

In the RA task, both %APCC and %APVC differed significantly between cognates and non-cognates [respectively, U = 139.32, z = 4.01, p < 0.001 and U = 134.71, z = 2.73, p = 0.006]. Again, the percentage of correct consonants and vowels was higher in the non-cognates than in the cognates.

Table 2. Aligned percentages of consonants correct (%APCC) and vowels correct (%APVC) tallied for cognates and non-cognates in the DR and RA conditions

	Cognates		Non-cognates	
	DR	RA	DR	RA
APCC%	52%	42%	65%	62%
APVC%	34%	18%	52%	37.5%

4.3. Vocabulary-Size Effects

In this section we analyze the effects of individual differences operationalized in terms of vocabulary-size (Nation 2006) on participants' vowel production accuracy. We calculated vowel and consonant accuracy scores for each participant, derived from the ratio between the absolute number of correct vowels/consonants and the absolute number of vowels/consonants attempted. Table 3 shows the vocabulary sizes and the accuracy scores for each individual in both the DR and the RA task; values closer to 1 indicate high accuracy, conversely, values closer to 0 indicate low production accuracy. Vowel accuracy scores exhibited relative variability across participants, but consonant accuracy scores reached a ceiling for virtually all the participants. Therefore, we made the decision to limit the analysis to the vowel sounds only because it is widely accepted that they have a greater weight in the intelligibility of L1-Spanish L2-English speech, as shown in recent works assessing training effects on L2 perception and production of English vowels (Carlet and Cebrian 2019; Fouz-González 2020). To test the effects of vocabulary-size on vowel accuracy ratios we ran two independent-samples t-tests. No significant differences were found between HV and LV participants for either the DR vowel accuracy ratios [t (21) = 1.40, p = 0.176] or the RA vowel accuracy ratios [t (21) = 1.33, p = 0.196], indicating that participants with high vocabulary-size did not necessarily produce the vowel sounds more accurately than participants with low vocabulary-size.

Table 3. Vocabulary-sizes and vowel and consonant accuracy scores for each individual participant[4]

Participant	Vocabulary-Size	Vowel Ratios DR	Consonant Ratios DR	Vowel Ratios RA	Consonant Ratios RA
AD	4,500	0.130	0.820	0.213	0.769
AG	3,600	0.572	0.886	0.271	0.732
AS	5,900*	0.900	0.900	0.789	0.938
CP	3,100	0.800	0.879	0.633	0.863
CB	1,700	0.305	0.767	0.183	0.804
CE	700	0.300	0.656	0.140	0.715
GG	3,300	0.591	0.851	0.625	0.919
JB	2,600	0.453	0.869	0.305	0.810
JP	3,000	0.892	0.891	0.669	0.852
JM	6,200*	0.766	0.899	0.626	0.875
KP	1,300	0.563	0.613	0.627	0.827
KG	1,800	0.771	0.850	0.593	0.801
KA	5,700*	0.663	0.923	0.471	0.882
MD	900	0.754	0.925	0.528	0.856
MM	1,200	0.762	0.913	0.603	0.826
MC	4,800	0.544	0.857	0.431	0.779
ND	7,000*	0.363	0.720	0.366	0.843
NB	2,900	0.858	0.960	0.658	0.868
NA	1,900	0.628	0.896	0.355	0.812
PC	3,700	0.892	0.913	0.641	0.885
PU	1,800	0.630	0.708	0.479	0.689
SM	2,500	0.491	0.777	0.330	0.840
ZC	2,800	0.347	0.832	0.223	0.820

4.4. Lexical Frequency and Word Familiarity

To test for the effect of word frequency on the percentage of vowels correctly produced we averaged the number of vowels produced correctly across speakers and separately for each task condition. The mean DR and RA ratios averaged across words and task conditions were not correlated with word frequency in either of the two elicitation conditions: RA [$n = 40$, $r = 0.21$, $p = 0.179$], DR [$n = 40$, $r = 0.112$, $p = 0.491$]. This indicates that more frequent words were not necessarily better produced than less frequent words.

[4] The scores correspond to the ratios between the total number of vowels/consonants correct and the total number of target vowels/consonants for each elicitation condition. Learners with vocabulary-sizes close to or above 6,000 words are marked with *.

5. Discussion

The present study explored the factors that influenced Spanish-Catalan intermediate EFL learners' pronunciation in an instructed setting. Specifically, we analyzed the effects of task type, cognate status of the target words, vocabulary-size and lexical frequency. Overall, the number of correct consonants was higher than for vowels, probably due to the closer match between the consonant inventories of Spanish/Catalan and English. In contrast, the number of correct vowels being much lower was likely because learners perceptually assimilate two or more English vowels to a single L1 category. Regardless of these differences, both task type and cognate status influenced learners' segmental production of English vowels and consonants. In the following subsections we discuss the effects of these variables in more detail.

5.1. Task and Cognate Effects

Overall, students performed significantly better in the DR condition than in the RA condition. Various reasons may account for these results. We could argue that the RA condition requires more attentional resources because learners have to map the graphemes to different linguistic units such as phonemes, syllables and/or morphemes. Although English, Spanish and Catalan share the same alphabet, there is a clear cross-linguistic difference in the degree of orthographic transparency between the three languages. English has an opaque orthography with little correspondence between sound and spelling. However, Spanish and Catalan are transparent languages with virtually a one-to-one correspondence between sound and letter. This means that, when faced with unfamiliar words, learners were not able to inhibit the rules of the L1 (perfect match between sounds and letters), thus causing a great number of pronunciation errors, especially in cognates. In contrast, the DR condition is an imitation task that relies on participants' working memory capacity and mimicry skills. One surprising finding is that even learners with small vocabulary sizes performed relatively well in this task suggesting that even if they were not familiar with some of the target words, they could produce them accurately.

As expected, less pronunciation errors were found in non-cognates relative to cognates. This is in line with previous studies examining production of L2 sounds in bilingual populations (Amengual 2012, 2016; Mora and Nadeu 2012; Rallo-Fabra 2015). Most learners transferred the Spanish pronunciation of cognate words, for instance *bilingual* was mostly produced as */bɪˈlɪŋgwal/ or */bɪˈlɪŋgwe/ and this was not only the case in the RA condition, in many cases the same error occurred in the DR condition, suggesting that learners had only one form stored in their mental lexicon and that it served for all three languages—Spanish, Catalan and English.

5.2. Does Vocabulary-Size Matter?

Vocabulary-size did not influence vowel production accuracy, which means that both groups of participants, HV and LV, performed in a similar fashion. These results are

surprising, since Bundgaard-Nielsen et al (2011a, 2011b) found that HV Japanese listeners categorized Australian English vowels more consistently than their LV peers, suggesting that lexical development enhances L2 phonological learning. Two factors may account for the discrepancy between our results and those of previous research. On the one hand, the threshold between HV and LV learners in the Bundgaard-Nielsen et al. studies was set at 6,000 words, based on the evidence that this is the minimum number of words needed to understand spoken English (Nation 2006). In our case only two participants had vocabulary sizes higher than 6,000 words, which might explain the absence of significant differences in vowel production by the two groups of participants. On the other hand, the effects reported by Bundgaard-Nielsen et al. were based on the results of perceptual assimilation studies and, therefore, they might not be extendable to L2 production.

The lack of significant differences between LV and HV learners in the present study could also be explained by a methodological limitation. The test administered to the participants was a receptive vocabulary test which included many cognate words. We speculate that some of the learners could have guessed the meaning of these cognates even if they had not heard them before or did not know how to pronounce them. Further to this, receptive vocabulary might increase if learners are exposed to massive written input, which does not necessarily imply that they know the meaning as well as the pronunciation of the new words. Future research examining L2 pronunciation will have to address this limitation and include aurally-elicited productive vocabulary tests.

5.3. Lexical Effects

We did not find any significant correlations between vowel production accuracy and lexical frequency. The finding that most participants did not have many difficulties producing the high-frequency words such as *people*, *rights* and *money*—even though they have incongruent orthographies—suggests that lexical frequency did play a role. The possibility exists that learners were exposed to accented renditions of the target words either by their peers or even by their teachers. It follows that if the learning environment cannot provide authentic input, learners will not be able to establish mental phonological representations for the new lexical items. Some of the errors found in the DR such as the substitutions of *brother* for *border*, *policía* for *politician* and *found* for *funds* suggest that word recognition might be a prerequisite for accurate speech production. These errors clearly illustrate that, when learners could not retrieve a given word from their mental lexicon, they would automatically elicit a phonological neighbor, understood as a "similar-sounding form-based representation that is activated in memory on the basis of stimulus input and compete for recognition" (Vitevith and Luce 2016, 73). It seems that the DR task involves a process of word recognition with two stages: a) activation of similar-sounding form-based representations in memory; and b) subsequent competition for recognition among these activated representations. In other words, learners might store the phonological form of a word along with the

lexical item. This would imply that words may be discrete entities just like phonemes, as has been suggested in exemplar models (Port 2017). Future research should address the role of lexical frequency using methods other than corpora and with populations at the early stages of L2 learning so that it is easier to control whether exposure to new lexical items occurs in the written form or in the auditory form. We tentatively speculate that if the auditory form precedes the written form learners will be more successful at achieving a more target-like pronunciation.

5.4. Pedagogical Implications

The data reported above reveal that for Spanish-Catalan learners to achieve less accented pronunciation of English as an L2, intervention methods that target specific segmental and/or suprasegmental aspects of pronunciation are required. Informal interviews with the school teachers whose students participated in this study revealed that no specific pre-planned pronunciation lessons were implemented to raise the students' awareness of the cross-linguistic pronunciation differences of the cognate words between Spanish/Catalan and English. Pronunciation does not receive much attention in the EFL classrooms in Spain and it is often restricted to listen-and-repeat drills or corrective feedback to tackle specific segmental errors. This scenario is in line with previous work on class observation in Francophone Canada (Foote et al. 2016), where analysis of video-recordings of class observations showed that pronunciation teaching only comprised 10% of language-related episodes. Interestingly, the video recordings also revealed large discrepancies between teachers' self-reports about time devoted to pronunciation teaching and the actual time spent teaching this skill in class.

Some participants in the current study acknowledged that words such as *journalist*, *politician*, *Jewish* or *disease* were new to them. The only exposure they had received was through watching a film with English subtitles. Recent work on the effects of watching original version films with captions by Catalan-Spanish learners has shown that it is necessary to insert specific questions to raise students' awareness so that their attention can be focused on pronunciation (Wisniewska and Mora 2020). If this is not done, it is very unusual that students pay attention to how specific words are pronounced. We must bear in mind that 90% of the participants had vocabulary sizes that did not reach 6,000 words. This means that they possibly focused all their attentional resources on understanding the meaning of vocabulary and did not pay much attention to form.

Future research avenues should attempt to bridge the gap between research and teaching practice. In-service teachers should familiarize themselves with specific pronunciation resources available open-access and gradually build the basis for a pronunciation teaching infrastructure. In turn, researchers should treat pedagogical implications as the central goal of their research, not as an afterthought (Levis 2019). Further studies investigating pronunciation by L1-Spanish intermediate learners have attempted to test the effectiveness of training methods to modify inexperienced learners' fossilized pronunciation. In a review of pronunciation instruction, Saito and Plonsky (2019) provide empirical evidence

that explicit teaching methods in the form of articulatory and/or auditory training have a positive impact on learners' pronunciation skills. In this line, Gómez-Lacabex and Gallardo-del-Puerto (2018) trained Spanish-Basque young learners to perceive the differences between accented and native-like instances of English schwa, that is */ˈbeɪkon/ vs. /ˈbeɪkən/. They found that both explicit and implicit teaching methodologies made a positive impact in terms of raising students' awareness on discerning between accented and native-like instances of this vowel. Ultimately, pronunciation instruction should encourage students to engage in meaningful communicative tasks, which involves "using the newly acquired knowledge (e.g. speaking language for meaning with correct pronunciation forms)" (Saito and Plonsky 2019, 693), so that L2/FL speech becomes automatized, fluent, error-free and effortless.[5]

Works Cited

Amengual, Mark. 2012. "Interlingual Influence in Bilingual Speech: Cognate Status Effects in a Continuum of Bilingualism." *Bilingualism: Language and Cognition* 15: 517-30.

—. 2016. "Cross-Linguistic Influence in the Bilingual Mental Lexicon: Evidence of Cognate Effects in the Phonetic Production and Processing of a Vowel Contrast." *Frontiers in Psychology* 7 (article 617): 9-25.

Bassetti, Bene and Nathan Atkinson. 2015. "Effects of Orthographic Forms on Pronunciation in Experienced Instructed Second Language Learners." *Applied Psycholinguistics* 36 (1): 67-91.

Best, Catherine T. and Michael D. Tyler. 2007. "Non-Native and Second Language Speech Perception: Commonalities and Complementarities." In Bohn and Munro 2007, 13-34.

Blecua, B. et al., eds. 2022. Propuestas en fonética experimental. Enfoques metodológicos y nuevas tecnologías. Girona: Universitat de Girona.

Bohn, Ocke-Schwen and Murray J. Munro, eds. 2007. *Language Experience in Second Language Speech Learning. In Honor of James Emil Flege*. Amsterdam and Philadelphia: John Benjamins.

Bundgaard-Nielsen, Rikke L., Catherine T. Best and Michael D. Tyler. 2011a. "Vocabulary-Size Matters: The Assimilation of Second-Language Australian English Vowels to First-Language Japanese Vowel Categories." *Applied Psycholinguistics* 32: 51-67.

—. 2011b. "Vocabulary-Size Is Associated with Second-Language Vowel Perception Performance in Adult Learners." *Studies in Second Language Acquisition* 33: 433-61.

[5] This research was funded by grants FFI2017-84479-P from the Spanish Ministry of Economy, Industry and Competitiveness and PID2020-117804GB-100 from the Spanish Ministry of Science and Innovation. I am indebted to Karen Lesley Jacob and Clara Vega Munar for their assistance with data collection and analysis.

Carlet, Angelica and Juli Cebrian. 2019. "Assessing the Effect of Perceptual Training on L2 Vowel Identification, Generalization and Long-term Effects." In Nyvad et al. 2019, 91-119.

Chang, Li-Yun, David C. Plaut and Charles A. Perfetti. 2016. "Visual Complexity in Orthographic Learning: Modeling Learning across Writing System Variations." *Scientific Studies of Reading* 20 (1): 64-85.

Derwing, Tracey M. and Murray J. Munro. 2015. *Pronunciation Fundamentals. Evidence-Based Perspectives for L2 Teaching and Research*. Amsterdam and Philadelphia: John Benjamins.

Erdener, Doğu. 2016. "Basic to Applied Research: The Benefits of Audio-Visual Speech Perception Research in Teaching Foreign Languages." The *Language Learning Journal* 44 (1): 124-32.

— and Denis K. Burnham. 2005. "The Role of Audio-Visual Speech and Orthographic Information in Non-native Speech Production." *Language Learning* 55 (2): 191-228.

Field, Anthony. 2012. *Discovering Statistics Using IBM SPSS Statistics*. London: Sage.

Foote, Jennifer Ann et al. 2016. "Pronunciation Teaching Practices in Communicative Second Language Classes." The *Language Learning Journal* 44 (2): 181-96.

Fouz-González, Jonás. 2020. "Using Apps for Pronunciation Training: An Empirical Evaluation of the English File Pronunciation App." *Language Learning and Technology* 24 (1): 62-85.

Gaillard, Stéphanie and Annie Tremblay. 2016. "Linguistic Proficiency Assessment in Second Language Acquisition Research: The Elicited Imitation Task." *Language Learning* 66 (2): 419-47.

Gallardo-del-puerto, Francisco, Esther Gómez-Lacabez and María Luisa García Lecumberri. 2009. "Testing the Effectiveness of Content and Language Integrated Learning in Foreign Language Contexts: The Assessment of English Pronunciation." In Ruiz de Zarobe and Jiménez-Catalán 2009, 63-80.

Gimenes, Manuel and Boris New. 2016. "Wordlex: Twitter and Blog Word Frequencies for 66 Languages." *Behaviour Research Methods* 48 (3): 963-72.

Gómez-Lacabex, Esther and Francisco Gallardo-del-Puerto. 2018. "Explicit Phonetic Instruction vs. Implicit Attention to Native Exposure: Phonological Awareness of English Schwa in CLIL." *International Review of Applied Linguistics in Language Teaching* 58 (4): 419-42.

Gorba, Celia et al. 2021. "Investigating the Role of Cognate Status in L2 Production and Perception." In Blecua et al. 2022, 122-26.

Hansen-Edwards, Jette G. and Mary Zampini, eds. 2008. *Phonology and Second Language Acquisition*. Amsterdam and Philadelphia: John Benjamins.

Hayes-Harb, Rachel and Shannon Barrios. 2021. "The Influence of Orthography in Second Language Phonological Acquisition." *Language Teaching* 54: 297-326.

Henderson, Alice et al. 2015. "The English Pronunciation Teaching in Europe Survey: Factors Inside and Outside the Classroom." In Mompean and Fouz-González 2015, 260-91.

LEVIS, John. 2005. "Changing Contexts and Shifting Paradigms in Pronunciation Teaching." *TESOL Quarterly* 39 (3): 369-77.

—. 2019. "Connecting the Dots between Pronunciation Research and Practice." Lecture given at the 6th International Conference on English Pronunciation: Issues and Practices, Skopje, May 2019.

— and Linda Grant. 2011. "Integrating Pronunciation into ESL/EFL Classrooms." *TESOL Journal* 12 (2): 13-19.

MAIRANO, Paolo and Fabian Santiago. 2020. "What Vocabulary Size Tells Us about Pronunciation Skills: Issues in Assessing L2 Learners." *Journal of French Language Studies* 30: 141-60.

MCALLISTER-BYUN, Tara and Yvan Rose. 2016. Analyzing Clinical Phonological Data Using Phon. *Seminars in Speech and Language* 37 (2): 85-105.

MOMPEAN, José Antonio and Jonás Fouz-González, eds. 2015. *Investigating English Pronunciation: Trends and Directions*. London: Palgrave-Macmillan.

MORA, Joan C. and Marianna Nadeu. 2102. "L2 Effects on the Perception and Production of a Native Vowel Contrast in Early Bilinguals." *International Journal of Bilingualism* 16: 484-500.

MUNRO, Murray J. 2008. "Foreign Accent and Speech Intelligibility." In Hansen-Edwards and Zampini 2008, 193-218.

NATION, Paul. 2006. "How Large a Vocabulary is Needed for Reading and Listening?" *Canadian Modern Language Review* 63 (1): 59-82.

—. 2010. *Test your Vocabulary Online with VocabularySize.com*. [Accessed October 18, 2015].

NYVAD, Anne Mette et al., eds. 2019. *A Sound Approach to Language Matters - In Honor of Ocke-Schwen Bohn*. Aarhus: Aarhus University.

OTWINOWSKA-KASZTELANIC, Agnieszka. 2009. "Raising Awareness of Cognate Vocabulary as a Strategy in Teaching English to Polish Adults." *Innovation in Language Learning and Teaching* 3 (2): 131-47.

PISKE, Thorsten, Ian R.A. MacKay and James E. Flege. 2001. "Factors Affecting Degree of Foreign Accent in an L2: A Review." *Journal of Phonetics* 29: 191-215.

PORT, Robert. 2007. "How are Words Stored in Memory? Beyond Phones and Phonemes." *New Ideas in Psychology* 25: 145-72.

RALLO-FABRA, Lucrecia. 2015. "Can Non-Native Speakers Reduce English Vowels in a Native-Like Fashion? Evidence from L1-Spanish L2-English Bilinguals." *Phonetica* 72: 162-81.

— and María Juan-Garau. 2011. "Assessing FL Pronunciation in a Semi-Immersion Setting: The Effects of CLIL Instruction on Spanish-Catalan Learners' Perceived Comprehensibility and Accentedness." *Poznan Studies in Contemporary Linguistics* 47: 96-108.

RINGBOM, Hakan. 2007. *Crosslinguistic Similarity in Foreign Language Learning*. Clevedon: Multilingual Matters.

Rose, Yvan and Brian MacWhinney, comp. 2015. *PHON Database System for the Study of Phonetics and Phonology* (version 2.1.4).

Ruiz de Zarobe, Yolanda and Rosa Maria Jiménez Catalán, eds. 2009. *Content and Language Integrated Learning. Evidence from Research in Europe*. Clevedon: Multilingual Matters.

Saito, Kazuya and Luke Plonsky. 2019. "Effects of Second Language Pronunciation Teaching Revisited: A Proposed Measurement Framework and Meta-Analysis." *Language Learning* 69: 652-708.

Schmitt, Norbert and Michael McCarthy, eds. 1997. *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge UP.

Trofimovich, Pavel et al. 2012. "A Frequency-Based Approach to L2 Phonological Learning: Teacher Input and Student Output in an Intensive ESL Context." *TESOL Quartely* 46: 176-86.

Vitevitch, Michael S. and Paul A. Luce. 2016. "Phonological Neighborhood Effects in Spoken Word Perception and Production." *Annual Review of Linguistics* 2: 7.1-7.20.

Vokic, Gabriela. 2011. "When Alphabets Collide: Alphabetic First-language Speakers' Approach to Speech Production in an Alphabetic Second Language. *Second Language Research* 27 (3): 391-417.

Wisniewska, Natalia and Joan C. Mora. 2020. "Can Captioned Video Benefit Second Language Pronunciation?" *Studies in Second Language Acquisition* 42 (3): 599-624.

Woore, Robert. 2018. "Learners' Pronunciations of Familiar and Unfamiliar French Words: What Can they Tell Us about Phonological Decoding in an L2?" The *Language Learning Journal* 46 (4): 456-69.

Received 29 May 2021 Revised version accepted 15 November 2021

Lucrecia Rallo Fabra holds a PhD in Linguistics from the Universitat de Barcelona. She is currently a senior lecturer in English Studies at the Universitat de les Illes Balears and teaches undergraduate and master's courses in English phonetics and phonology, applied linguistics and research methods. She has authored articles in *Journal of Phonetics* and *Phonetica* and has also contributed book chapters in research monographs such as *Second Language Speech Learning: Theoretical and Empirical Progress* (2021).