

Article

The Interplay of Syllable Structure and Consonant Sonority in L2 Speech Segmentation

Juan José Garrido-Pozú 

Department of Modern Languages and Literatures, Furman University, Greenville, SC 29613, USA;
juan.garrido@furman.edu

Abstract: The present study investigated whether L1 and L2 Spanish speakers show sensitivity to matching/mismatching syllable structure and consonant sonority in lexical segmentation in Spanish. A total of 81 English–Spanish learners and 72 Spanish–English learners completed a fragment-monitoring task. They listened to lists of Spanish words as they saw a CV or CVC syllable (e.g., “pa” or “pal”) and pressed a button when the word began with the syllable shown on the screen. The task manipulated syllable structure (CV or CVC) and consonant sonority (fricative, nasal, or liquid) of target syllables and carrier words. Target syllables either matched or did not match the structure of the first syllable in target carrier words (e.g., “pa—pa.lo.ma”; “pa—pal.me.ra”). The results showed that consonant sonority modulated sensitivity to syllable structure in both groups of participants. Spanish–English learners responded faster to matching syllable structure in words that had a fricative or a nasal as the second consonant, and English–Spanish learners responded faster only with a fricative consonant. Higher L2 Spanish proficiency correlated with faster target-syllable identification, but sensitivity to matching/mismatching structure did not vary as a function of proficiency. The study highlights the influence of phonetic factors in the development of L2 lexical segmentation routines.

Keywords: speech segmentation; perception; consonant sonority; syllable; second-language acquisition



Citation: Garrido-Pozú, Juan José. 2024. The Interplay of Syllable Structure and Consonant Sonority in L2 Speech Segmentation. *Languages* 9: 103. <https://doi.org/10.3390/languages9030103>

Academic Editor: Joseph V. Casillas

Received: 12 December 2023

Revised: 4 March 2024

Accepted: 8 March 2024

Published: 18 March 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction and Background

A crucial part of the listening process entails identifying meaningful linguistic units in a continuous speech signal. Lexical identification in continuous speech is a complex task because sounds in the speech signal overlap with each other in time and space, even across word boundaries, and word boundaries are rarely marked explicitly in the acoustic signal. During the early stages of speech processing, listeners can parse the speech signal into prelexical phonological units that can be used to access the lexicon and activate lexical items (see [Floccia et al. 2012](#); [Pallier et al. 2001](#); [McQueen et al. 2006](#)). Although some accounts propose that the prelexical phonological units into which the speech stream is parsed are syllable-sized perceptual units (see [Savin and Bever 1970](#)), the specific nature of these units is yet to be determined. Guided by cross-linguistic evidence, many models of lexical processing include the availability of multiple perceptual units that can be used during prelexical processing (e.g., [Dupoux 1993](#); [Gaskell and Marslen-Wilson 1997](#); [Goldinger and Azuma 2003](#); [McClelland and Elman 1986](#)).

Cross-linguistic evidence suggests that among other factors, phonetic and phonological differences across languages may influence the types of perceptual routines in which listeners engage. Segmentation routines may vary based on the type of rhythm of a given language ([Ramus et al. 1999](#)). Native speakers of languages commonly classified as “syllable-timed” (i.e., languages whose rhythmic structure revolves around the syllable and whose syllables have a roughly equivalent duration; see [Conlen 2016](#); [Liu and Takeda 2021](#)), like Spanish, Italian, French, and Portuguese, use a syllable-like prelexical unit to parse the speech signal, and they take advantage of syllabic information for lexical segmentation (e.g., [Floccia et al. 2012](#); [Mehler et al. 1981](#); [Sebastián-Galles et al. 1992](#)). On the other hand,

native speakers of languages classified as “stress-timed” (i.e., languages in which stressed syllables are spread out consistently with equal amounts of time in between; see [Conlen 2016](#); [Liu and Takeda 2021](#)), like English, Dutch, and German, rely on patterns of lexical stress for speech segmentation during prelexical processing (e.g., [Cutler et al. 1986](#); [Dupoux 1993](#); [Cutler and Norris 1988](#); [Mehler et al. 1981](#)).

Although available studies on lexical segmentation report differences in the segmentation routines of speakers of syllable-timed and stress-timed languages, there is compelling evidence that argues against this rhythm class typology based on timing distinctions (for reviews, see [Arvaniti and Rodriquez 2013](#); [Fletcher 2010](#); [Loukina et al. 2011](#)). Studies have found evidence against the view that timing is the base of speech rhythm and instead suggest that factors such as speaking rate and F_0 play a more consistent role than timing in discrimination between languages (see [Arvaniti and Rodriquez 2013](#)). The reported differences in segmentation of languages like Spanish in comparison to languages like English may well be due to timing distinctions, differences in speaking rate, or differences in the complexity and predictability of their syllable patterns, which could facilitate using syllable structure as a cue for segmentation in some languages but not in others. Regardless of the nature of the distinction, studies have found evidence of syllabic segmentation in languages commonly classified as syllable-timed but not in languages commonly classified as stress-timed (see [Cutler et al. 1986](#); [Dupoux 1993](#); [Mehler et al. 1981](#)). The present study uses this classification based on timing to build upon prior evidence in segmentation studies.

It is still unclear how bilinguals and second language (L2) learners who speak a syllable-timed language and a stress-timed language carry out lexical segmentation and what prelexical units they use to achieve it. The present study investigated lexical segmentation in L1 speakers and L2 learners of Spanish through a syllable detection paradigm ([Mehler et al. 1981](#)). The study tested whether L2 learners of Spanish exhibit evidence of the use of segmentation routines that are motivated only by the phonological structure of their stress-timed L1 (English), or whether they develop speech segmentation routines that are specific to the phonological structure of their syllable-timed L2 (Spanish). The current study extends previous research to examine the influence of phonetic/phonological factors such as syllable structure and consonant sonority in prelexical processing.

Early models of lexical processing assumed that the syllable had a central role in speech recognition and that syllable-like units mediated the mapping of the speech signal onto the lexicon. For example, the Standard Syllabic Model ([Mehler et al. 1981](#)) and the Cascade Syllabic Model ([Dupoux 1993](#)) used a restrictive approach to speech processing which was centered around the syllable. Their assumptions were mostly based on evidence from syllable-timed languages like Italian, Spanish, and French (e.g., [Mehler et al. 1981](#)), but they lacked support from stress-timed languages like English (e.g., [Bradley et al. 1993](#); [Cutler et al. 1986](#)). A more dynamic model that accounts for cross-linguistic differences is the Semi-Syllables Model ([Dupoux 1993](#)), which assumes different perceptual units depending on the phonological properties of each language. For example, Spanish speakers may use syllable-like units, whereas English speakers may use feet or other stress-based units. Evidence of the availability of different perceptual units is found in studies reporting that segmentation routines vary across languages based on their phonological composition, and that listeners can develop language-specific segmentation routines (e.g., [Bradley et al. 1993](#); [Cutler et al. 1992](#); [Detey and Nespoulous 2008](#); [Katayama 2015](#)). In the present study, the Semi-Syllables Model ([Dupoux 1993](#)) can account for differences in the segmentation routines of Spanish L1 and English L1 subjects and the presence of language-specific segmentation.

More recent models of speech processing do not include syllabic units as universal perceptual units. They allow for the availability of different perceptual units across languages and processing levels (e.g., [Luce et al. 2000](#); [Marslen-Wilson and Welsh 1978](#); [Marslen-Wilson 1987](#); [Norris 1994](#); [Shook and Marian 2013](#)). While most models assume a prelexical level of representation, the perceptual units at this level differ in nature. Importantly, a few models explicitly incorporate the role of syllable structure in the mapping of the speech

signal onto the lexicon (e.g., [Luce et al. 2000](#); [Shook and Marian 2013](#)). Although the syllable is not a central unit in these models, syllable structure helps modulate lexical activation and access to different degrees.

There is abundant evidence of the use of syllabic information in speech segmentation in native speakers of syllable-timed languages (e.g., French: [Cutler et al. 1986](#); [Mehler et al. 1981](#); Catalan: [Sebastián-Galles et al. 1992](#); Spanish: [Bradley et al. 1993](#); Portuguese: [Morais et al. 1989](#); Italian: [Flocchia et al. 2012](#); see [Simonet 2019](#) for a review). Many of these studies reported that under specific conditions, monolingual speakers exhibit a target-type-by-word-type interaction in syllable detection in monitoring tasks. In a monitoring task, participants listen to lists of words as they see a specific target syllable or fragment on the screen. They are asked to detect the words that begin with the syllable or fragment that they see on the screen. A target-type-by-word-type interaction is observed when participants respond faster, or more accurately, to targets that match the syllabic structure of the first syllable of the carrier word than to targets that do not match the structure of the first syllable of the carrier word. For example, French monolinguals respond faster to the fragment “ba” in the word “balance” than in the word “balcon”, and they respond faster to the fragment “bal” in the word “balcon” than in the word “balance” ([Mehler et al. 1981](#)).

This sensitivity to matching/mismatching syllabic information in the target-type-by-word-type interaction has been interpreted as evidence of syllabic segmentation during prelexical processing. However, studies have failed to replicate this interaction in stress-timed languages like Dutch and English (e.g., English: [Cutler et al. 1986](#); Dutch: [Zwitserslood 1989](#)), suggesting that speech segmentation is language-specific and that syllabic segmentation occurs only in speakers of syllable-timed languages, while speakers of stress-timed languages use non-syllabic segmentation routines. These results are more consistent with the premises of the Semi-Syllables Model ([Dupoux 1993](#)), which assumes different perceptual units across different languages.

Syllabic effects in lexical segmentation emerged not only in studies using monitoring tasks but also in studies using word-spotting tasks ([Dumay et al. 2002](#); [Kang and Nam 2014](#); [Garrido-Pozú 2023](#)), cross-modal priming tasks ([Tabossi et al. 2000](#); [Tagliapietra et al. 2009](#)), and syllable-reversal and partial-repetition tasks ([Content et al. 2001](#)). Effects of syllable structure are not restricted to lexical segmentation only. Differences in L1–L2 syllable structure have also been found to modulate word recognition and production, causing a facilitation effect when both the L1 and the L2 are stress-timed languages, but not when the L1 is syllable-timed and the L2 is stress-timed ([Martínez García 2021](#); [Martínez García and Tremblay 2015](#)). Cross-linguistic differences in patterns of syllable structure also affect L2 consonant perception and production (e.g., [Cheng and Zhang 2015](#); [Yasufuku and Doyle 2021](#)) and L2 auditory word learning (e.g., [Hamada and Goya 2015](#)), showing that syllabic effects are observed in different dimensions.

Considering that monolingual speakers employ segmentation routines that are specific to the phonology of their specific language, it is then relevant to examine whether L2 learners and bilinguals employ the same segmentation routines for both of their languages or develop different language-specific routines for each of their languages. Bilinguals present an interesting scenario because they deal with two different linguistic systems, each of which has its own phonological structure with different rhythmic patterns. Studying L2 segmentation routines allows us to investigate to what extent L2 segmentation routines are restricted by the L1, whether it is possible to develop different language-specific segmentation routines, and what factors influence the development of L2 segmentation routines, among other matters that remain unclear.

Studies addressing bilingual speech segmentation in syllable-timed languages are scarce, and those available have provided mixed findings. Available evidence suggests that L2 segmentation is constrained by the phonology of the L1 or the dominant language ([Cutler et al. 1989, 1992](#)). For instance, [Cutler et al. \(1992\)](#) tested French–English and English–French early bilinguals in three tasks: two fragment-monitoring tasks (in English and French) and a word-spotting task in English. The results of the French fragment-

monitoring task revealed that only the French-dominant early bilinguals exhibited the target-type-by-word-type interaction, which is typically observed in French monolinguals. However, results of the English fragment-monitoring task yielded no target-type-by-word-type interaction for any of the groups. The results of the word-spotting task in English indicated that only the English-dominant early bilinguals used lexical stress to segment the speech signal, which is commonly observed in English monolinguals. [Cutler et al. \(1992\)](#) showed that rhythm-based language-specific segmentation routines may be mutually exclusive. The early bilinguals in this study exhibited evidence of the segmentation routine that was motivated by their dominant language only. French-dominant participants behaved like French monolinguals and exhibited syllable-based segmentation; and English-dominant participants behaved like English monolinguals. However, the lack of syllabic effects for French-dominant bilinguals with English words shows that these bilinguals employed two different segmentation routines, a syllable-based routine for French and a non-syllable-based routine for English. Thus, bilinguals with a syllable-timed L1 can develop unmarked non-syllabic segmentation routines for L2 perception whereas bilinguals with a stress-timed L1 seem to not develop syllabic routines for an L2 ([Cutler et al. 1986, 1989](#)), which highlights the role of language dominance and L1 type in the development of L2 segmentation strategies.

Further evidence from Spanish–English bilinguals reveals differences in the segmentation routines employed by Spanish-dominant bilinguals and Spanish monolinguals. [Bradley et al. \(1993\)](#) tested Spanish monolinguals and English monolinguals in two fragment-monitoring tasks, one in Spanish and one in English. Additionally, they also tested Spanish–English early bilinguals (Spanish-dominant) in a fragment-monitoring task in Spanish. The results showed no target-type-by-word-type interaction for English monolinguals in any of the tasks. On the other hand, a target-type-by-word-type interaction was observed with Spanish monolinguals in the Spanish task but not in the English task. Spanish–English bilinguals did not exhibit a target-type-by-word-type interaction with Spanish words. These results contrast with previous evidence indicating that French monolinguals use syllable-based segmentation even for English words ([Cutler et al. 1986](#)), suggesting that syllabic routines in Spanish monolinguals might not be as stable as in French monolinguals because they only exhibit syllabic influence in Spanish and not in English. In addition, the lack of syllabic effects in the Spanish–English bilinguals in [Bradley et al. \(1993\)](#) conflicts with the results for the French–English bilinguals in [Cutler et al. \(1992\)](#). These bilinguals did not exhibit the syllabic effects typically attributed to speakers with a syllable-timed L1. Importantly, the bilinguals in [Bradley et al. \(1993\)](#) had been living in an English-speaking country for an extended period of time. The difference between the performance of Spanish monolinguals and Spanish–English bilinguals with Spanish stimuli could indicate that either syllabic segmentation in Spanish is unstable and could easily be abandoned, or acquisition of a stress-timed L2, coupled with extended immersion in the L2 context, may cause listeners to modify their approach to input representation and abandon syllabic routines even for L1 materials.

In addition to syllable structure, phonetic information may affect speech segmentation as well. The Syllable Onset Segmentation Hypothesis (henceforth SOSH; [Content et al. 2000](#)) suggests that syllabic information aids speech segmentation by providing possible points of alignment for the detection of possible word onsets. Since locating word boundaries implies establishing syllable boundaries, SOSH claims that speech segmentation is affected by syllable structure. Syllable onsets constitute reliable points of alignment for the lexical search process because syllable onsets often coincide with word onsets and are more salient in the signal than offsets. Consequently, detecting syllable onsets and detecting syllable offsets in the speech signal involve different processes and constraints. Syllable onset detection is more reliable and effective than syllable offset detection. In fact, in word-spotting tasks, listeners identify target words faster and more accurately when the onset aligns with a syllable onset than when the offset aligns with a syllable offset (see [Dumay et al. 2002](#)). On the other hand, syllable offset detection is influenced by the level of sonority

of intervocalic consonants (Content et al. 2001). A common scale for sonority includes vowels > approximants (glides and liquids) > nasals > fricatives > affricates > stops, with vowels being the most sonorous sounds and stops being the least sonorous. According to SOSH, more sonorous intervocalic consonants are more likely to be assigned as codas (offset of the previous syllable), while less sonorous consonants are more likely to be assigned as onsets. Thus, L2 listeners are more likely to place a syllable boundary before less sonorous intervocalic consonants and after more sonorous intervocalic consonants.

Intervocalic consonant sonority typically affects segmentation when listeners are engaged in detection of a syllable offset, but not in detection of a syllable onset. Monolingual speakers of Spanish commonly assign single intervocalic consonants as onsets rather than offsets. Based on SOSH, English-speaking L2 learners of Spanish are more likely to exhibit native-like Spanish segmentation with less sonorous intervocalic consonants than with more sonorous intervocalic consonants. In other words, one can predict that English-speaking L2 learners of Spanish are more likely to assign fricative intervocalic consonants (e.g., /s/ in “basura”) as onsets of the following syllable (ba.su.ra) while liquid intervocalic consonants (e.g., /l/ in “balance”) are more likely to be assigned as offsets of the previous syllable (bal.an.ce). The present study manipulates syllable structure and intervocalic consonant sonority to further study patterns in segmentation of L2 Spanish in English-speaking L2 learners of Spanish and Spanish-speaking L2 learners of English. In addition, the present study includes L2 learners of different levels of proficiency to examine how segmentation routines develop as L2 acquisition progresses.

The Present Study

The present study investigates lexical segmentation in L1 and L2 Spanish. Available evidence of syllabic effects in segmentation of Spanish comes from studies addressing early bilinguals (e.g., Bradley et al. 1993), but studies on segmentation in adult L2 learners of Spanish are scarce. This study examines lexical segmentation in adult Spanish-speaking L2 learners of English and English-speaking L2 learners of Spanish. Studying these two groups of L2 learners provides an opportunity to assess whether listeners with a stress-timed L1 (English) develop language-specific routines for segmentation of a syllable-timed L2 (Spanish) and whether listeners with a syllable-timed L1 (Spanish) modify their segmentation routines with acquisition of a stress-timed L2 (English). In addition, while studies using word-spotting tasks have highlighted the role of consonant sonority in segmentation of French (Dumay et al. 2002), the role of consonant sonority in segmentation of Spanish is still unclear. The current study also explores the influence of phonetic/phonological factors such as intervocalic consonant sonority and syllable structure in lexical segmentation of Spanish to test how acoustic-phonetic information modulates syllable detection and syllabic effects. Notably, available studies on syllabic effects in segmentation have not addressed possible effects of language proficiency in segmentation routines. The present study includes subjects with different levels of proficiency to examine how L2 segmentation routines vary with higher L2 proficiency. This study is driven by the following research questions:

- RQ1: Are L2 learners sensitive to matching/mismatching syllable structure during a fragment-monitoring task in Spanish?
- RQ2: Does intervocalic consonant sonority modulate L2 learners’ sensitivity to matching/mismatching syllable structure during a fragment-monitoring task in Spanish?
- RQ3: Does L2 proficiency modulate L2 learners’ sensitivity to syllable structure during a fragment-monitoring task in Spanish?

The first research question examines whether Spanish–English learners and English–Spanish learners are sensitive to syllable structure during syllable detection in Spanish. Previous studies suggest that segmentation routines in bilinguals are constrained by the phonological composition of their first or dominant language (Carroll 2004; Cutler et al. 1989, 1992). Specifically, speakers with a stress-timed L1 do not typically develop L2 syllable-based segmentation routines, and it is unclear whether acquisition of an L2 is related to a

modification of segmentation routines even for L1 materials. Following previous studies, the present study predicted that Spanish–English learners are likely to show sensitivity to matching/mismatching syllable structure, but English–Spanish learners are not. If syllabic effects are observed in the English–Spanish group, the results will serve as evidence that L2 learners can develop two separate language-specific segmentation routines, and that with L2 acquisition, English–Spanish learners can develop a segmentation routine that is not motivated by their L1.

The second research question asks whether intervocalic consonant sonority modulates L2 learners' sensitivity to syllable structure in syllable detection in Spanish. SOSH (Content et al. 2000) predicts that the sonority of intervocalic consonants modulates whether they are assigned as onsets or offsets, thus affecting syllabification and segmentation. Evidence consistent with this claim has been observed with French monolinguals (Dumay et al. 2002), but evidence of sonority effects in L2 and bilingual segmentation is scarce. Based on SOSH and previous studies, the present study predicted that English–Spanish learners are more likely to exhibit syllabic effects, which resemble the effects in Spanish monolinguals, with less sonorous intervocalic consonants.

Finally, the third research question explores the role of L2 proficiency in syllable detection in Spanish. It is still unclear whether L2 proficiency modulates the development of L2-specific segmentation routines and whether the effects of consonant sonority and syllable structure in segmentation vary as a function of L2 proficiency. The present study predicted that higher proficiency in Spanish results in more sensitivity to matching/mismatching syllable structure.

2. Materials and Methods

2.1. Participants

A total of 153 participants were recruited: 81 adult English-speaking L2 learners of Spanish (henceforth English L1–Spanish L2) and 72 adult Spanish-speaking L2 learners of English (henceforth Spanish L1–English L2). The English L1–Spanish L2 subjects were native speakers of English who were born and raised in the United States and had no intensive exposure to L2 Spanish before puberty. These subjects were recruited from Spanish classes in a large university in the northeast of the US. They received either course credit or monetary compensation for their participation in the study. The Spanish L1–English L2 subjects were native speakers of Spanish who were born and raised in Peru and had no intensive exposure to L2 English before puberty. They were recruited from a large English language institute in Lima, Peru, and received monetary compensation for their participation in the study. All participants completed a background questionnaire in their L1 where they provided biographical and linguistic information.

All participants were between 18 and 40 years old. For English L1–Spanish L2 subjects, the mean age was 22.2 ($SD = 5.06$), and for Spanish L1–English L2 subjects, the mean age was 31.4 ($SD = 6.76$). In total, 49 participants identified as male, and 79 participants identified as female. Overall, the English L1–Spanish L2 subjects started taking L2 Spanish classes at school at age 11.9 ($SD = 6.33$), studied L2 Spanish for about 8.51 years ($SD = 3.82$), were exposed to L2 Spanish about 18% of their time per week ($SD = 15.3$), and spent about 14.4% ($SD = 15.3$) of their overall time producing language either writing or speaking in L2 Spanish. On the other hand, on average, the Spanish L1–English L2 subjects started taking L2 English classes at age 15.2 ($SD = 5.12$), studied L2 English for 4.67 years ($SD = 4.03$), were exposed to L2 English about 31.6% of their time per week ($SD = 20.7$), and spent about 27.6% ($SD = 22.8$) of their overall time producing language either writing or speaking in L2 English. Participants included L2 speakers with a wide range of L2 proficiency levels.

2.2. Tasks and Procedures

2.2.1. Proficiency Tests

Language proficiency was operationalized using four proficiency tests, two in Spanish and two in English. Two tests measured proficiency based on vocabulary size, and two tests

measured grammatical knowledge. For English proficiency, participants completed the Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer and Broersma 2012). The LexTALE provided a measure of participants' vocabulary size in English through a visual lexical decision task in which participants judged 60 lexical items (40 words and 20 pseudowords). They were required to read the items one by one and indicate whether each item was an existing word in the English language by pressing a key. Scores are based only on accuracy, and they ranged from 0 to 100. This task took approximately 5 min. In addition to the LexTALE, participants also completed a Cloze Test in English, designed by Brown (1980) and later adapted by Martínez García (2016). The test measures English proficiency based on grammatical knowledge. Participants read a passage about the evolution and progress of humans that contained 50 multiple-choice blanks and indicated which words best completed the text. The blanks corresponded to content words and function words. Scores ranged from 0 to 50, and the test took approximately 15 min.

For Spanish proficiency, participants completed the Spanish version of the LexTALE (LexTALE-ESP; Izura et al. 2014). The LexTALE-ESP uses a similar methodology and design as the LexTALE. Participants completed a visual lexical decision task in which they judged 90 lexical items (60 words and 30 pseudowords) and indicated whether each item was a real word in Spanish with a key press. Scores ranged from −20 to 60 and are based on accuracy. Typically, native speakers score above 50 in the LexTALE-ESP. The task took about 5 min to complete. Additionally, participants completed an abridged version of the “*Diploma de Español como Lengua Extranjera*” (DELE), which is an official accreditation degree of fluency in the Spanish language issued and recognized by the Ministry of Education, Culture, and Sports of Spain (Sagarra and Herschensohn 2010; Sagarra et al. 2024). The abridged version of the DELE measured grammatical knowledge and reading comprehension in Spanish via a multiple-choice test. It contained 60 question items (21 for reading comprehension and 39 for grammar). Scores ranged from 0 to 56, and the test took about 20 min.

Proficiency scores were used as continuous variables in the analysis. Scores of the two tests in each language were highly correlated. Figure 1 shows the distribution of proficiency scores in the LexTALE and the Cloze Test for both groups of participants. Figure 2 plots the proficiency scores in the LexTALE-ESP and the DELE for both groups of participants. In Figures 1 and 2, all the scores are standardized as z-scores.

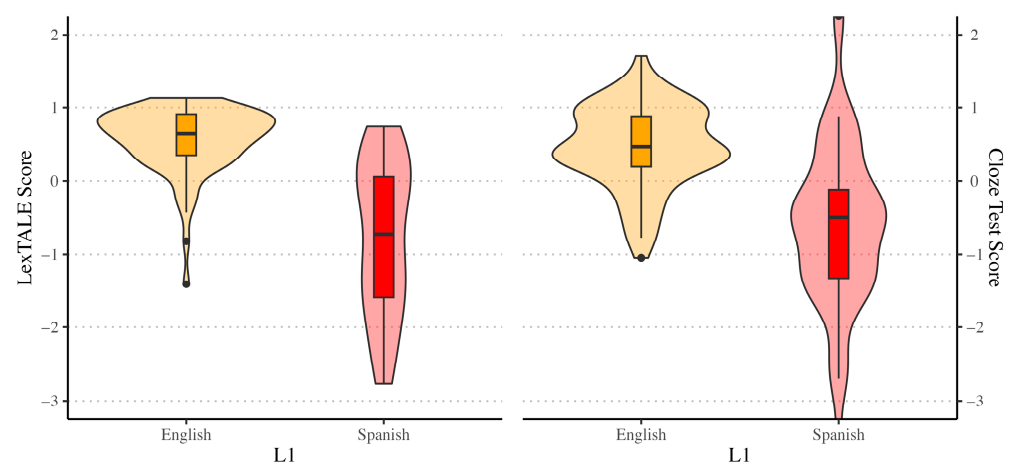


Figure 1. English proficiency scores of both groups of participants, English L1–Spanish L2 subjects and Spanish L1–English L2 subjects. The left panel shows scores of the LexTALE, and the right panel shows scores of the Cloze Test. Scores are displayed as z-scores.

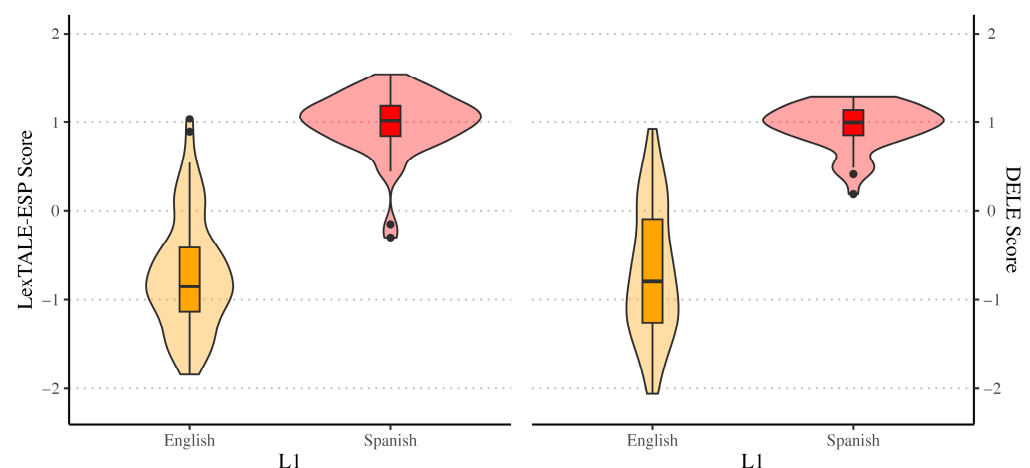


Figure 2. Spanish proficiency scores of both groups of participants, English L1–Spanish L2 subjects and Spanish L1–English L2 subjects. The left panel shows scores of the LexTALE-ESP, and the right panel shows scores of the DELE. Scores are displayed as z-scores.

2.2.2. Fragment-Monitoring Task

The fragment-monitoring task (henceforth FMT) tested whether participants were sensitive to syllable structure and consonant sonority during fragment detection in segmentation of Spanish. The task aimed to replicate the target-type-by-word-type interaction found in previous studies in Spanish monolinguals (Bradley et al. 1993), French monolinguals (Cutler et al. 1986, 1992; Mehler et al. 1981), and Portuguese monolinguals (Morais et al. 1989). This FMT followed the design and procedure of the task used in Bradley et al. (1993).

Participants listened to lists of isolated words in Spanish (e.g., “paloma”, “palmera”) as they saw one fragment on the screen (e.g., “pa”). Participants were instructed to press a button only when the word they heard began with the fragment shown on the screen. They were asked to respond as fast as possible. In all the trials, participants permanently saw the instructions at the top center of the screen, which said “Presiona la barra de espacio solo si la palabra empieza con la siguiente secuencia” (Press the space bar only if the word begins with the following sequence). The target fragment was displayed in the center of the screen, and 500 ms later, a list of four words was presented aurally. The target fragment changed after presentation of four words. Only one of the four words in each list contained the target fragment. The words that matched the target fragment varied in the structure of the first syllable. For example, both “paloma” and “palmera” have “pa” at the beginning, but “pa” matches the structure of the first syllable of “pa.lo.ma” and does not match the first syllable of “pal.me.ra”. Participants were expected to take longer to identify “pa” in “palmera” than in “paloma” because of the mismatch in syllable structure between the fragment and the target word. The FMT recorded response times (RTs) and accuracy (correct and incorrect identifications). RTs were measured from the ending of the stimuli words to prevent any possible effects of differences in the duration of the segments.

The participants listened to a total of 48 experimental words, which were trisyllabic nouns with stress on the penultimate syllable. The experimental items varied in the structure of the first syllable. Half of the words began with a CV first syllable (e.g., CV—*pa.lo.ma*), and the other half began with a CVC first syllable (e.g., CVC—*pal.me.ra*). Each CV word shared the same three initial phonemes with a CVC word (e.g., “paloma” and “palmera”), so the total list contained 24 CV.C–CVC pairs. Experimental items also varied in the level of sonority of the second consonant, which could be a liquid (e.g., “paloma” and “palmera”), a nasal (e.g., “sonido” and “sonrisa”), or a fricative (e.g., “basura” and “bastillo”). In the carrier words that had a liquid second consonant, these were realized as voiced alveolar laterals in CV words and as voiced dental or alveolar laterals in CVC words due to coarticulation. Nasals were realized as voiced alveolar nasals in CV words

and as voiced dental or alveolar nasals in CVC words due to coarticulation. Fricatives were realized as voiceless alveolar fricatives in both CV and CVC words. In CVC words, fricatives were followed by a voiceless dental plosive or a voiceless velar plosive. Visual target fragments were CV (e.g., “pa”) or CVC (e.g., “pal”) sequences. Each of the 48 target words was placed within a randomized list with three other filler and distractor words. The fillers were nouns, verbs, or adjectives. No fillers included the CV or CVC target fragment. To make sure that participants base their responses on the detection of complete target sequences, distractors were designed as catch trials. Some distractors shared a segment or two of the target sequence. A set of 3 practice lists were included at the beginning of the task.

The stimuli were recorded by a male native speaker of Peruvian Spanish. During the recording session, the speaker read each word from a computer screen embedded at the beginning of the sentence “____ es la palabra correcta” (____ is the correct word). The stimuli were recorded using a Shure SM58 microphone and a Marantz Solid State Recorder PMD670, at a sampling rate of 44.1 kHz and 16-bit quantization. Each item was recorded three times, and the best recording was selected manually based on clarity. The speaker was instructed to read the items at a normal rate, and he did not have knowledge of the target words to ensure that no emphasis was given to target words.

The FMT was designed in two versions. Each version included the same experimental words, but the target fragments were manipulated across versions. Since the goal of the task was to compare how long participants took to identify CV and CVC target fragments like “pa” and “pal” in CV words like “paloma” vs. CVC words like “palmera”, each experimental item had to be presented with a CV and a CVC target fragment. To avoid one participant encountering the same experimental item twice, in each version, the target fragments corresponded to the first syllable of half of the target words only. The rest of the target sequences were shorter or longer than the first syllable of the other half of the words. For example, in Version A, participants encountered “pa” with “paloma” and “pal” with “palmera”, and in Version B, participants encountered “pal” with “paloma” and “pa” with “palmera”. Across both versions, both types of target fragments were combined with each of the carriers. Each participant completed only one version of the task, and the distribution was counterbalanced. A list of experimental items is provided in Table S1 in the Supplementary Materials.

2.2.3. Data Analysis

The data were analyzed using mixed-effects models. Since the goal of the present study was to analyze the effects of matching/mismatching syllable structure (match type), sonority, and Spanish proficiency in syllable detection in each group, the data of the English L1 subjects and the Spanish L1 subjects were submitted to separate models. English L1 subjects’ RTs were analyzed as a function of the fixed effects *match type* (match or mismatch), *sonority* (fricative, nasal, or liquid), and *proficiency*. Spanish L1 subjects’ RTs were analyzed as a function of the fixed effects *match type* (match or mismatch) and *sonority* (fricative, nasal, or liquid). The models included higher-order interactions between all the fixed effects. Only RTs of correct responses were included in the analysis. The random effects’ structure included by participant and by item random intercepts. Main effects and interactions were assessed by partitioning the variance hierarchically via nested model comparisons. The models were best fit when including the random effects. Alpha was set at 0.05. The statistical analyses were carried out using R (R Core Team 2022). The analysis used the packages *lme4* (Bates et al. 2015) and *lmerTest* (Kuznetsova et al. 2017) to fit the mixed-effects models and *emmeans* (Lenth 2022) for multiple comparisons.

3. Results

This section presents the results of the Spanish L1–English L2 subjects followed by the results of the English L1–Spanish L2 subjects. Spanish L1 subjects identified the target fragment correctly (i.e., they identified the word that contained the target sequence

correctly) in over 90% of the instances. The analysis of RTs of correct responses revealed a main effect of *match type* ($\chi^2(1) = 17.51, p < 0.001$) and a significant interaction between *match type* and *sonority* ($\chi^2(2) = 6.04, p = 0.04$). Overall, Spanish L1 subjects responded to the match condition (e.g., *pa—paloma*) about 31 ms faster than to the mismatch condition (e.g., *pal—paloma*) ($SE = \pm 0.008$). However, this effect was modulated by consonant sonority. Pairwise comparisons revealed that Spanish L1 subjects responded faster to the match condition when the intervocalic consonant of the carrier word was a fricative ($p = 0.02$) or a nasal ($p = 0.003$). With a fricative intervocalic consonant, Spanish L1 subjects responded to the match condition approximately 42 ms faster than to the mismatch condition ($SE = \pm 0.013$). With a nasal intervocalic consonant, Spanish L1 subjects responded to the match condition approximately 48 ms faster ($SE = \pm 0.013$). Figure 3 shows mean RTs in milliseconds of Spanish L1 subjects as a function of *match type* and *sonority*. RTs in this plot are aggregated across both target fragment structures and word types (CV or CVC). Figure 4 shows mean RTs in milliseconds as a function of *match type*, *sonority*, and *item structure* (CV words and CVC words). Figure 4 shows that Spanish L1 subjects responded faster to matching syllable structure regardless of consonant sonority when the target word had a CV first syllable (e.g., *pa—pa.lo.ma*). When the target word had a CVC first syllable (e.g., *pal.me.ra*), Spanish L1 subjects benefited from matching syllable structure with a fricative or a nasal consonant but not with a liquid consonant. A summary of the final model used to make inferences on the data of Spanish L1 subjects can be found in Table 1.

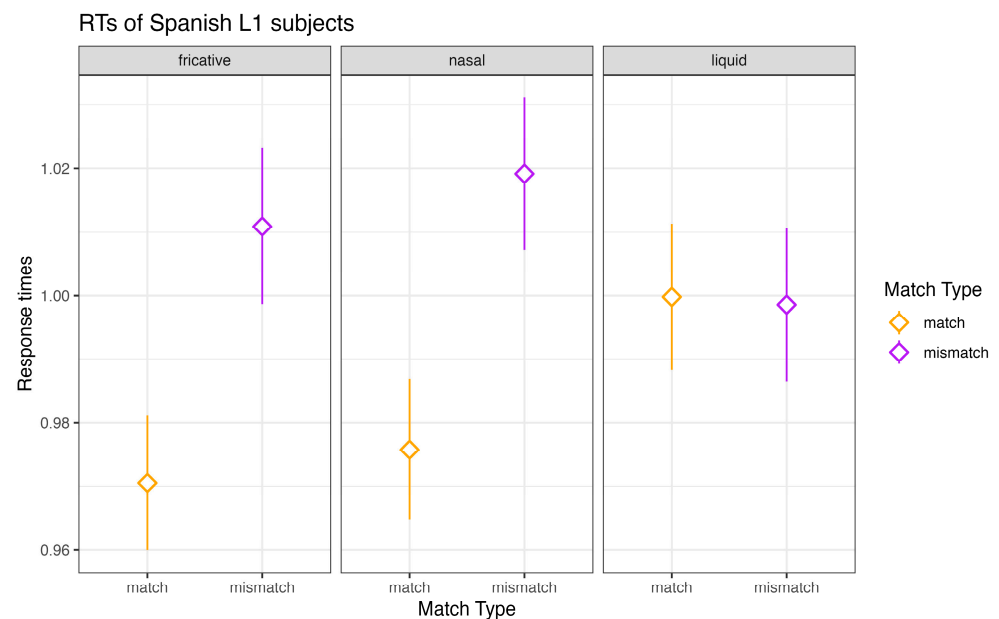


Figure 3. Response times in milliseconds of Spanish L1–English L2 subjects as a function of *match type* and *sonority*. RTs are aggregated across target types and word types.

Table 1. Summary of the final model used to make inferences on the data of Spanish L1 subjects.

	Estimate	Std. Error	df	t Value	Pr(> t)
(Intercept)	0.97	0.02	103.45	39.86	<0.001
match_typemismatch	0.04	0.01	2344.05	3.16	0.002
sonorityliquid	0.03	0.02	63.93	1.15	0.25
sonoritynasal	0.003	0.02	64.37	0.16	0.88
match_typemismatch:sonorityliquid	−0.04	0.02	2343.58	−1.92	0.05
match_typemismatch:sonoritynasal	0.01	0.02	2343.82	0.36	0.72

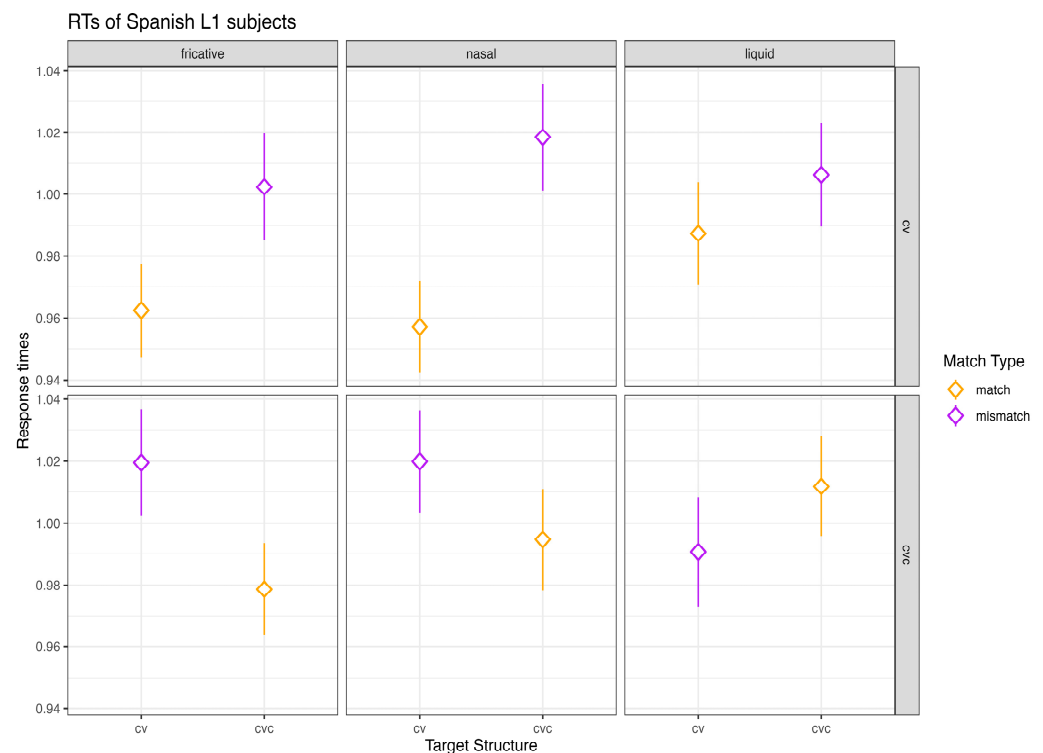


Figure 4. Response times in milliseconds of Spanish L1–English L2 subjects as a function of *target structure*, *match type*, *sonority*, and *item structure*. The top panel shows RTs for CV words, and the bottom panel shows RTs for CVC words. The match condition in the top panel represents a CV target fragment (e.g., “pa”) and a CV word (e.g., “paloma”). In the bottom panel, the match condition represents a CVC target fragment (e.g., “pal”) and a CVC word (e.g., “palmera”).

Regarding English L1–Spanish L2 subjects, they also exhibited high accuracy rates, detecting target fragments correctly in over 90% of the instances overall. The analysis of RTs of correct responses revealed no main effects of *match type*, *sonority*, or *proficiency*. However, the models revealed a significant interaction between *match type* and *sonority* ($\chi^2(2) = 6.12$, $p = 0.04$). Pairwise comparisons revealed that English L1 subjects responded faster to the match condition only when the intervocalic consonant of the carrier word was a fricative ($p = 0.04$). With a fricative intervocalic consonant, they responded to the match condition approximately 37 ms faster than to the mismatch condition ($SE = \pm 0.013$) for both CV and CVC words. Interestingly, although not statistically significant, there was a trend for English L1 subjects to respond about 25 ms faster to the mismatch condition than to the match condition in CV words with a liquid intervocalic consonant and to a lesser degree with a nasal intervocalic consonant (see Figure 5). Spanish L2 proficiency did not yield any significant effects on the response times of English L1 subjects. However, there was a trend for overall RTs to decrease as proficiency increased. Post hoc analyses yielded a significant negative correlation between overall RTs and L2 proficiency, but the correlation was not strong ($r(3383) = -0.22$, $p < 0.001$). For words with a fricative intervocalic consonant, English L1 subjects seemed to benefit more from matching syllable structure with higher L2 proficiency. Figure 5 shows mean RTs in milliseconds of English L1 subjects as a function of *match type* and *sonority* with RTs aggregated across target and word type. Figure 6 shows mean RTs as a function of *match type*, *sonority*, and *item structure* (CV words and CVC words). Figure 6 shows a shift in the pattern of RTs in CV words across levels of consonant sonority, with matching syllable structure facilitating target detection in words with a fricative consonant and mismatching structure facilitating target detection in words with a nasal or liquid consonant. A summary of the final model used to make inferences on the data of English L1 subjects can be found in Table 2.

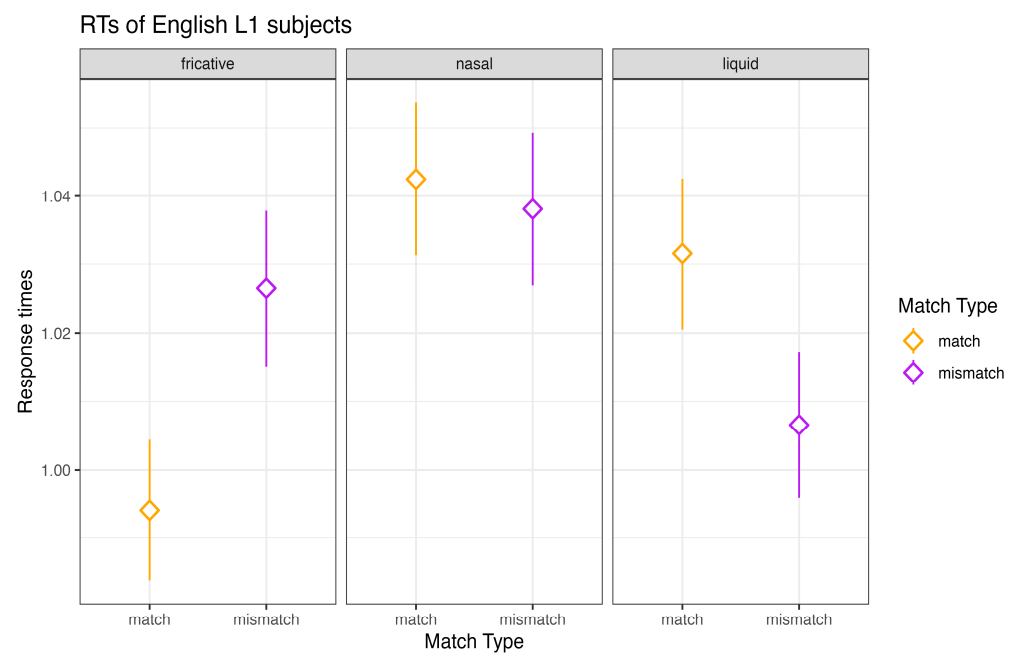


Figure 5. Response times in milliseconds of English L1–Spanish L2 subjects as a function of *match type* and *sonority*. RTs are aggregated across target types and word types.

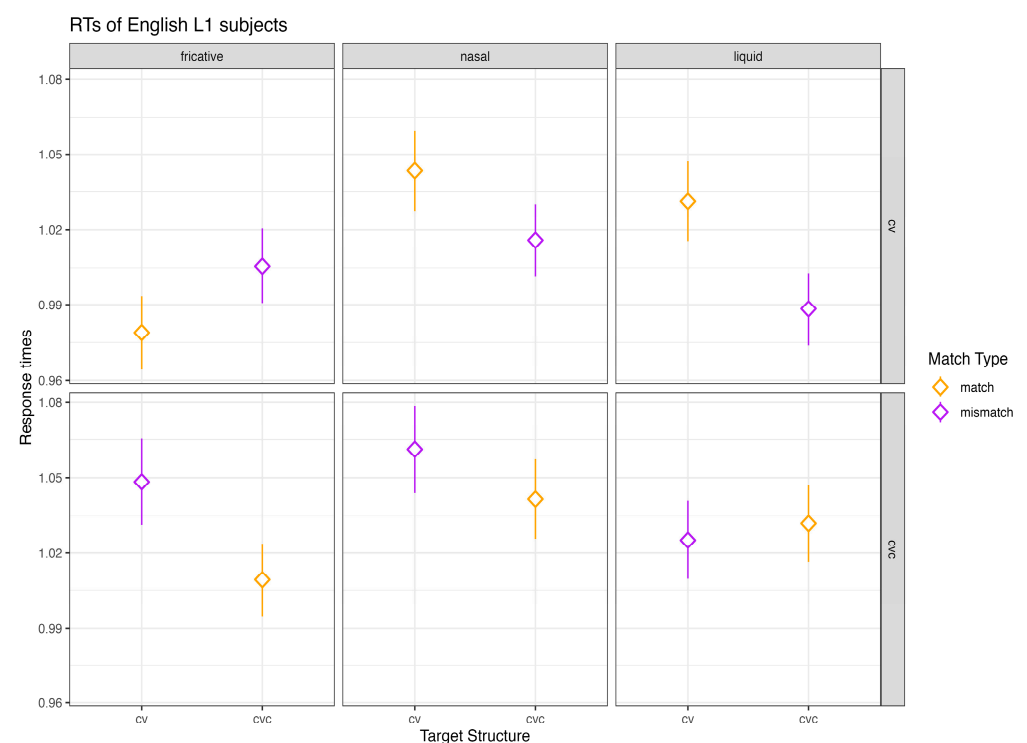


Figure 6. Response times in milliseconds of English L1–Spanish L2 subjects as a function of *target structure*, *match type*, *sonority*, and *item structure*. The top panel shows RTs for CV words, and the bottom panel shows RTs for CVC words. The match condition in the top panel represents a CV target fragment (e.g., “pa”) and a CV word (e.g., “paloma”). In the bottom panel, the match condition represents a CVC target fragment (e.g., “pal”) and a CVC word (e.g., “palmera”).

Table 2. Summary of the final model used to make inferences on the data of English L1 subjects.

	Estimate	Std. Error	df	t Value	Pr(> t)
(Intercept)	0.99	0.02	127.88	41.37	<0.001
match_typedismatch	0.04	0.01	3101.34	2.86	0.004
sonorityliquid	0.03	0.02	64.47	1.51	0.14
sonoritynasal	0.04	0.02	64.47	1.99	0.05
match_typedismatch:sonorityliquid	−0.05	0.02	3122.08	−2.46	0.01
match_typedismatch:sonoritynasal	−0.03	0.02	3114.66	−1.38	0.17

4. Discussion

The present study investigated lexical segmentation of Spanish in English-speaking L2 learners of Spanish and Spanish-speaking L2 learners of English. More specifically, the study examined whether L2 learners exhibit the use of language-specific segmentation when monitoring syllabic fragments in Spanish. The study used a fragment-monitoring task to test whether L2 learners were sensitive to matching/mismatching syllable structure and assess the influence of intervocalic consonant sonority and L2 proficiency on lexical segmentation of Spanish.

The first research question focused on L2 learners' sensitivity to matching/mismatching syllable structure during segmentation of Spanish. Previous studies showed that segmentation routines in bilinguals are constrained by the phonological composition of their first or dominant language, with English L1 speakers having more difficulty to develop L2-specific syllable-based segmentation routines (Bradley et al. 1993; Carroll 2004; Cutler et al. 1989, 1992; Dupoux 1993). The results of the current study are in line with previous findings since Spanish L1 subjects and English L1 subjects benefited from matching syllable structure in different ways. Participants identified target syllables faster when they matched the structure of the first syllable of the carrier word, but this effect of matching/mismatching syllable structure was modulated by consonant sonority, affecting RTs of Spanish L1 subjects and English L1 subjects differently. When including all target and word types, Spanish L1 subjects responded faster to the match condition with a fricative and a nasal intervocalic consonant, but not with a liquid; and English L1 subjects responded faster to the match condition only with a fricative intervocalic consonant. When looking at the interplay of target type, word type, and consonant sonority, Spanish L1 subjects responded faster to matching syllable structure across all levels of sonority when the carrier word had a CV first syllable. English L1 subjects, on the other hand, responded faster to matching syllable structure with CV words with a fricative consonant, and they responded faster to mismatching syllable structure with CV words with a nasal or liquid consonant. Both groups exhibited different patterns which corresponds to the L1/L2 status of Spanish and the type of segmentation strategies that are motivated by participants' L1.

The presence of syllabic effects in the Spanish L1–English L2 subjects in the present study is consistent with previous studies that reported syllabic effects in L1 speakers of languages with predictable syllable patterns like Spanish (e.g., Bradley et al. 1993), French (e.g., Cutler et al. 1989, 1992; Dumay et al. 2002; Mehler et al. 1981), Portuguese (e.g., Morais et al. 1989), and Italian (e.g., Floccia et al. 2012). However, these results differ from the results of Bradley et al. (1993), where Spanish-dominant early bilinguals did not exhibit syllabic effects in Spanish. The discrepancy may correspond to differences in the age and mode of acquisition and L2 immersion since the participants of the present study acquired L2 English after puberty through formal instruction, while those in Bradley et al. (1993) lived immersed in the context of the L2 since early childhood, which may have influenced the segmentation strategies that they employed. Importantly, the presence of syllabic effects in English L1–Spanish L2 subjects in the present study differs from previous studies claiming that L1 speakers of stress-based languages like English are unable to develop syllable-based segmentation routines for an L2 (e.g., Carroll 2004; Cutler et al. 1989, 1992). In the present study, differences in syllabic structure between the target and the carrier word

affected RTs of English L1 subjects, but the pattern varied based on consonant sonority, highlighting the role of phonetic information in speech segmentation.

The second research question focused on the role of consonant sonority in speech segmentation. The Syllable Onset Segmentation Hypothesis (SOSH; [Content et al. 2000](#)) predicts that the sonority of intervocalic consonants modulates whether they are assigned as onsets or offsets. Less sonorous intervocalic consonants are syllabified into the next syllable as onsets (e.g., “pa.sa.je”), and more sonorous intervocalic consonants are more likely to be syllabified into the previous syllable as codas (e.g., “ba.l.an.ce”), thus affecting syllabification and segmentation. Evidence consistent with this claim has been observed with French monolinguals ([Dumay et al. 2002](#)), but evidence in L2 learners and bilinguals is scarce. The results of the current study are in line with the assumptions of SOSH. Intervocalic consonant sonority modulated the effects of matching/mismatching syllable structure in the same direction as predicted by SOSH. Spanish L1 subjects exhibited stronger syllabic effects overall with the less sonorous consonants, and English L1 subjects benefited from matching syllable structure only with the least sonorous consonant. As the level of sonority increased, English L1 subjects displayed a reversed effect, which can be explained by the influence of sonority in syllabification assumed by SOSH. English L1 subjects may have employed different syllabification for words with more sonorous intervocalic consonants, which does not match syllabification patterns expected for Spanish, causing the reversed effect. If, according to SOSH, English L1 subjects syllabified more sonorous consonants (liquids and nasals) as offsets of the previous syllable (e.g., “pal.o.ma” and “con.e.jo”), the mismatch condition in the present study actually represented a match for English L1 subjects (e.g., “pal—pal.o.ma”). Consequently, what the present study coded as mismatching structure actually facilitated target detection due to English L1 subjects applying different syllabification to words with more sonorous consonants.

The last research question asked about the role of L2 proficiency in segmentation of Spanish. In this study, L2 proficiency did not yield any significant effects or interactions. However, there was a trend for English L1–Spanish L2 subjects to detect target sequences faster as L2 proficiency increased, but their sensitivity to matching/mismatching information did not vary as much as a function of L2 proficiency. In general, there was a trend for English L1–Spanish L2 subjects to respond faster to the match condition than to the mismatch condition with fricative intervocalic consonants, and as proficiency increased, the difference in RTs for match/mismatch became bigger. This trend could point to the possibility of higher L2 proficiency increasing sensitivity to syllable structure, but further research is needed. Additionally, in the case of liquid intervocalic consonants, the results showed that English–Spanish learners responded faster to the mismatch condition at lower levels of proficiency, but as proficiency increased, there was a slight shift in the opposite direction to observe faster responses to the match condition at higher levels of proficiency. A similar pattern was observed for nasal intervocalic consonants but to a lesser degree. The trend could suggest that the effects of consonant sonority predicted by SOSH, especially those that do not match the syllabification typically used in L2 Spanish (with more sonorous consonants), can shift with higher L2 proficiency and begin to resemble patterns observed in L1 speakers of Spanish. Future research is needed to disentangle the influence of L2 proficiency in L2 segmentation.

The results of the present study support syllabic models of speech segmentation such as the Standard Syllabic Model ([Dupoux 1993](#); [Mehler et al. 1981](#)), the Cascade Syllabic Model ([Dupoux 1993](#)), and the Semi-Syllables Model ([Dupoux 1993](#)) with regard to the influence of syllable structure in speech segmentation. The present study does not claim that the syllable has a privileged role in speech perception, as some of these models state, but the results suggest that syllable structure is involved in the process of speech segmentation even for L2 learners with a stress-timed L1. Likewise, although the results of the present study do not provide word recognition data, they could also inform models of spoken-word recognition. The Cohort Model ([Marslen-Wilson and Tyler 1980](#); [Marslen-Wilson and Welsh 1978](#); [Marslen-Wilson 1987](#)) assumes that during lexical access, only lexical items

that coincide with the onset of the incoming input are co-activated as a cohort. If syllable structure is involved in this process, the Cohort Model could account for the results of the present study because subjects responded faster when the target syllable coincided with the structure of the first syllable of the incoming carrier word. Based on the Cohort Model, one could assume that in a fragment-monitoring task, listeners more easily activate lexical items that are in the cohort pre-activated by the target syllable shown on the screen. Similarly, the results of this study are also in line with the PARSYN model (Luce et al. 2000), which is an instantiation of the Neighborhood Activation Model (NAM) (Luce and Pisoni 1998). PARSYN assumes that lexical activation flows in a bottom-up fashion and that at lower levels of recognition, pre-lexical information such as syllable structure affects the activation of lexical items, with mismatching information being inhibited. These models, however, do not account for the interactions between syllable structure and consonant sonority observed in segmentation and syllabification in the present study. Future models should take into account the influence of phonetic and phonological factors within and across languages to explain the processes involved in speech segmentation in bilingual and multilingual individuals.

5. Conclusions

The present study showed that matching/mismatching syllable structure between a syllabic fragment and a target word affects lexical segmentation in Spanish-speaking L2 learners of English and English-speaking L2 learners of Spanish, but intervocalic consonant sonority modulates this effect. Importantly, syllable structure and consonant sonority affected Spanish L1–English L2 subjects and English L1–Spanish L2 subjects differently, which showed that segmentation routines are constrained by the composition of a learner's L1. The study provides evidence that L2 learners with a stress-based L1 can exhibit syllabic effects when segmenting a syllable-based L2, which suggests that they are able to develop language-specific routines that are not motivated by the phonological composition of their L1. The study informs models of speech segmentation regarding the influence of syllable structure and the role of prelexical phonological and phonetic information in speech segmentation.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/languages9030103/s1>, Table S1: Experimental stimuli.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Rutgers University (protocol code Pro2021000015 and date of approval 3/2/2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are openly available in Open Science Framework at <https://osf.io/hgv83>.

Conflicts of Interest: The author declares no conflicts of interest.

References

- Arvaniti, Amalia, and Tara Rodriguez. 2013. The role of rhythm class, speaking rate, and F_0 in language discrimination. *Laboratory Phonology* 4: 7–38. [CrossRef]
- Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]
- Bradley, Dianne, Rosa Sánchez-Casas, and José García-Albea. 1993. The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes* 8: 197–233. [CrossRef]
- Brown, James. 1980. Relative merits of four methods for scoring cloze tests. *The Modern Language Journal* 64: 311–17. [CrossRef]
- Carroll, Susanne. 2004. Segmentation: Learning how to 'hear words' in the L2 speech stream. *Transactions of the Philological Society* 102: 227–54. [CrossRef]

- Cheng, Bing, and Yang Zhang. 2015. Syllable structure universals and native language interference in second language perception and production: Positional asymmetry and perceptual links to accentedness. *Frontiers in Psychology* 6: 1–17. [CrossRef]
- Conlen, Madeline. 2016. A Linguistic Comparison: Stress-Timed and Syllable-Timed Languages and Their Impact on Second Language Acquisition. Bachelor's thesis, Wayne State University, Detroit, MI, USA.
- Content, Alain, Nicolas Dumay, and Uli Frauenfelder. 2000. The role of syllable structure in lexical segmentation: Helping listeners avoid mondegreens. In *Spoken Word Access Processes*. Edited by Anne Cutler and James McQueen. Nijmegen: Max-Planck Institute for Psycholinguistics, pp. 39–42.
- Content, Alain, Ruth Kearns, and Uli Frauenfelder. 2001. Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language* 45: 177–99. [CrossRef]
- Cutler, Anne, and Dennis Norris. 1988. The Role of Strong Syllables in Segmentation for Lexical Access. *Journal of Experimental Psychology: Human Perception and Performance* 14: 113–21. [CrossRef]
- Cutler, Anne, Jacques Mehler, Dennis Norris, and Juan Segui. 1986. The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* 25: 385–400. [CrossRef]
- Cutler, Anne, Jacques Mehler, Dennis Norris, and Juan Segui. 1989. Limits on bilingualism. *Nature* 340: 229–30. [CrossRef]
- Cutler, Anne, Jacques Mehler, Dennis Norris, and Juan Segui. 1992. The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology* 24: 381–410. [CrossRef]
- Detey, Sylvain, and Jean-Luc Nespoulous. 2008. Can orthography influence second language syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters. *Lingua* 118: 66–81. [CrossRef]
- Dumay, Nicolas, Uli Frauenfelder, and Alain Content. 2002. The role of the syllable in lexical segmentation in French: Word-spotting data. *Brain and Language* 81: 144–61. [CrossRef]
- Dupoux, Emmanuel. 1993. The time course of prelexical processing: The syllabic hypothesis revisited. In *Cognitive Models of Speech Processing: The Second Sperlonga Meeting*. Edited by Gerry Altmann and Richard Shillcock. Hove: Lawrence Erlbaum Associates, Publishers, pp. 81–114.
- Fletcher, Janet. 2010. The prosody of speech: Timing and rhythm. In *The Handbook of Phonetic Sciences*. Edited by William Hardcastle, John Laver and Fiona E. Gibbon. Oxford: Wiley-Blackwell, pp. 521–602.
- Floccia, Caroline, Jeremy Goslin, José Junça De Morais, and Régine Kolinsky. 2012. Syllable effects in a fragment-detection task in Italian listeners. *Frontiers in Psychology* 3: 1–12. [CrossRef]
- Garrido-Pozú, Juan José. 2023. L2 speech segmentation for word recognition: The role of lexical stress and syllable structure. In *Proceedings of the 20th International Congress of Phonetic Sciences*. Edited by Radek Skarnitzl and Jan Volín. Prague: Guarant International, pp. 540–44.
- Gaskell, Gareth, and William Marslen-Wilson. 1997. Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes* 12: 613–56. [CrossRef]
- Goldinger, Stephen, and Tamiko Azuma. 2003. Puzzle-solving science: The quixotic quest for units in speech perception. *Journal of Phonetics* 31: 305–20. [CrossRef]
- Hamada, Megumi, and Hideki Goya. 2015. Influence of syllable structure on L2 auditory word learning. *Journal of Psycholinguistic Research* 44: 141–57. [CrossRef] [PubMed]
- Izura, Cristina, Fernando Cuetos, and Marc Brysbaert. 2014. Lextale-esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica* 35: 49–66.
- Kang, Jinwon, and Kichun Nam. 2014. The effects of syllable boundary and context on word recognition in Korean continuous speech. *International Journal of Intelligent Information and Database Systems* 8: 162–73. [CrossRef]
- Katayama, Tamami. 2015. Effect of phonotactic constraints on second language speech processing. *I-Perception* 6: 1–13.
- Kuznetsova, Alexandra, Per Brockhoff, and Rune Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82: 1–26. [CrossRef]
- Lemhöfer, Kristin, and Mirjam Broersma. 2012. Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods* 44: 325–43. [CrossRef]
- Lenth, Russell. 2022. Emmeans: Estimated Marginal Means, aka Least-Squares Means. Available online: <https://CRAN.R-project.org/package=emmeans> (accessed on 15 October 2023).
- Liu, Sha, and Kaye Takeda. 2021. Mora-timed, stress-timed, and syllable-timed rhythm classes: Clues in English speech production by bilingual speakers. *Acta Linguistica Academica* 68: 350–69. [CrossRef]
- Loukina, Anastassia, Greg Kochanski, Burton Rosner, Elinor Keane, and Chilin Shih. 2011. Rhythm measures and dimensions of durational variation in speech. *Journal of the Acoustical Society of America* 129: 3258–70. [CrossRef]
- Luce, Paul, and David Pisoni. 1998. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* 19: 1–36. [CrossRef]
- Luce, Paul, Stephen Goldinger, Edward Auer, and Michael Vitevitch. 2000. Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics* 62: 615–25. [CrossRef]
- Marslen-Wilson, William. 1987. Functional parallelism in spoken word-recognition. *Cognition* 25: 71–102. [CrossRef]
- Marslen-Wilson, William, and Alan Welsh. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10: 29–63. [CrossRef]

- Marslen-Wilson, William, and Lorraine Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 8: 1–71. [CrossRef]
- Martínez García, M. T. 2016. Tracking Bilingual Activation in the Processing and Production of Spanish Stress. Ph.D. thesis, University of Kansas, Lawrence, KS, USA.
- Martínez García, Maria Teresa. 2021. Syllable structure effects in word recognition by Spanish- and German-speaking second language learners of English. *Journal of the Spanish Association of Anglo-American Studies* 43: 1–21. [CrossRef]
- Martínez García, Maria Teresa, and Annie Tremblay. 2015. Syllable structure affects second-language spoken word recognition and production. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: The University of Glasgow. ISBN 978-0-85261-941-4. Paper number 0824. Available online: <http://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0824.pdf> (accessed on 15 October 2023).
- McClelland, James, and Jeffrey Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18: 1–86. [CrossRef] [PubMed]
- McQueen, James, Anne Cutler, and Dennis Norris. 2006. Phonological abstraction in the mental lexicon. *Cognitive Science* 30: 1113–26. [CrossRef]
- Mehler, Jacques, Jean Yves Dommergues, Uli Frauenfelder, and Juan Segui. 1981. The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior* 20: 298–305. [CrossRef]
- Morais, José, Alain Content, Luz Cary, Jacques Mehler, and Juan Segui. 1989. Syllabic segmentation and literacy. *Language and Cognitive Processes* 4: 57–67. [CrossRef]
- Norris, Dennis. 1994. Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52: 189–234. [CrossRef]
- Pallier, Christophe, Angels Colomé, and Núria Sebastián-Gallés. 2001. The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12: 445–49. [CrossRef]
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 15 October 2023).
- Ramus, Franck, Marina Nespor, and Jacques Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265–292. [CrossRef]
- Sagarra, Nuria, and Julia Herschensohn. 2010. The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua* 20: 2022–39. [CrossRef]
- Sagarra, Nuria, Laura Fernández-Arroyo, Cristina Lozano-Argüelles, and Joseph Casillas. 2024. Unraveling the complexities of second language lexical stress processing: The impact of first language transfer, second language proficiency, and exposure. *Language Learning*, 1–32. [CrossRef]
- Savin, Harris, and Thomas Bever. 1970. The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior* 9: 295–302. [CrossRef]
- Sebastián-Gallés, Nuria, Juan Segui, Emmanuel Dupoux, and Jacques Mehler. 1992. Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language* 31: 18–32. [CrossRef]
- Shook, Anthony, and Viorica Marian. 2013. The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism* 16: 304–324. [CrossRef] [PubMed]
- Simonet, Miquel. 2019. Phonological encoding and the phonology of Spanish—The role of the syllable. In *The Routledge Handbook of Spanish Phonology*, 1st ed. Edited by Sonia Colina, Fernando Martínez-Gil, Manel Lacorte and Javier Muñoz-Basols. London: Routledge.
- Tabossi, Patrizia, Simona Collina, Michela Mazzetti, and Marina Zopello. 2000. Syllables in the processing of spoken Italian. *Journal of Experimental Psychology: Human Perception and Performance* 26: 758–75. [CrossRef] [PubMed]
- Tagliapietra, Lara, Rachele Fanari, Simona Collina, and Patrizia Tabossi. 2009. Syllabic effects in Italian lexical access. *Journal of Psycholinguistic Research* 38: 511–26. [CrossRef] [PubMed]
- Yasufuku, Kanako, and Gabriel Doyle. 2021. Echoes of L1 syllable structure in L2 phoneme recognition. *Frontiers in Psychology* 12: 1–15. [CrossRef]
- Zwitslood, Pienie. 1989. The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32: 25–64. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.