4-2018

Stress Perception in L1 and L2 Spanish and English

Jessica Campbell

Stress Perception in L1 and L2 Spanish and English

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Arts in Linguistics from
The College of William and Mary

by

Jessica Amy Campbell

Accepted for Honors

Anya Lunden, Director

Leslie Cochrane

Kaitlyn Harrigan

Jonathan Arries

Williamsburg, VA
April 26, 2018

1. Introduction

Stress, as a linguistic phenomenon, is used nearly universally across language. In English, the words *America* and *Massachusetts* do not just differ in phonemes, but in suprasegmental properties, as well. That is, ***A****merica* has primary stress on the antepenult, and *Massa****chu****setts* has primary stress on the penult.

Word-level stress differs from intonational, or focus, stress in that it occurs in every word, regardless of whether or not the word is in focus. Some languages use word-level stress[1] contrastively, disambiguating words by placing emphasis on one syllable instead of another. Even if a language has stress-based minimal pairs, it can still have systematical assignment of secondary stress. Often languages that do not have a set syllable for stress have a "window" of syllables where the primary stress can lie. For example, Creek has a final two-syllable window (Martin, 2011).

Stress may be aligned left ("leftward") or aligned right ("rightward"), meaning that the primary stress is on that side of each word. For example, Koromfe stresses the initial syllable, and is therefore leftward (Rennison, 1997). Some languages have a binary stress pattern (nearly half of Lunden & Kalivoda's Stress Correlate Database), stressing a particular syllable and then stressing (usually secondarily) every other syllable to the left or right. Stress can be in either direction depending on the language. For instance, Czech stresses the initial syllable and every other syllable to the right, making it have leftward stress (because the primary stress starts on the left), whereas Fijian stresses the penultimate syllable and every other syllable to the left, making it have rightward stress (Dubeda & Votrubec, 2005; Dixon, 1988). Not all secondary stress directions are connected to the location of the primary stress. For example, Maquiritari has rightward stress but assigns secondary stress from left to right (Hall, 1988).

Moras are a measure of a syllable's "weight"; light syllables contain one mora, and heavy syllables contain more than one. A typical CV syllable has one mora, and languages vary in their qualifications for what else constitutes a mora. For example, diphthongs contain two moras, and some languages classify syllables with a coda consonant as containing two moras, as well. If a language bases stress on moras, moras are counted instead of syllables. Therefore, the location of stress may be affected by the presence of a coda, a diphthong, or other factors.

[1] Word-level stress will now be referred to as simply "stress".

English and Spanish are both rightward stress languages, with the primary stress in the final three-syllable window (Goedemans and van der Hulst, 2017; Bradley & Mackenzie. 2004). Though sources differ in the exact percentages of words in English and Spanish with penultimate and antepenultimate stress, Clopper's 2002 corpora-based study found that English four-syllable words with stress on the second syllable (6.77% of the corpus were four-syllable words with antepenultimate stress) occurred almost exactly as often as four-syllable words with stress on the third syllable (6.72% of the corpus were four-syllable words with penultimate stress). Four-syllable words fit into the 10% of Spanish words that are not two or three syllables, and penultimate stress occurs 64%-80% of the time (LaCross et al. 2016). While Dogil and Williams (1999) state that secondary stress is present on every other syllable preceding the primary stress in Spanish, they also find that "[secondary] stress in Spanish is often levelled out in speech" (p. 306). In contrast, English often does have perceptive secondary stress on every other syllable preceding. For example, in *Missis**si**ppi*, the penultimate syllable is stressed, but the first syllable has secondary stress. In *cami**se**ta* 'shirt', the penultimate syllable is stressed, but the first syllable does not have perceivable correlates for the secondary stress. This difference between the languages could cause native English speakers learning Spanish to erroneously pronounce Spanish words with secondary stress.

Spanish contains a large number of contrastively stressed words, meaning that some words differ only in terms of stress placement. For example, many verb conjugations rely solely on stress placement to distinguish themselves: *can**tó*** 'she sang' is very different from ***can**to* 'I sing'. These Spanish words are grammatical sentences in Spanish and do not require subjects, which would disambiguate the verb. A non-native speaker of Spanish, therefore, would have to perceive this stress as a native speaker would in order to glean meaning from their conversation partner's utterance. This perception, however, relies on a set of correlates that serve to cue the listener to the stress of the syllable. The following question and answer pairs demonstrate a sentence in which the contrastive stress occurs in a word not in focus. Note that there is a difference between the stress correlates in words that are in focus in a sentence and therefore contain special features based on sentence intonation; the words in question in this research are assumed not to be in focus.

(1) *¿**Can**to para mi mamá? No, **can**to para mi hermana.*

'I sing for my mom? No, I sing for my sister.'

(2) *¿Can**tó** para mi mamá? No, can**tó** para mi hermana.*

'She sang for my mom? No, she sang for my sister.'

A story teller using first person present point of view to tell a story could employ rhetorical questions such as (1) above. Such an instance is unusual, but it shows the possibility of this situation. In the case above, *hermana* 'sister' is the focus of both second sentences; the verb is not. Therefore, in these sentences, the stressed syllable of each verb (again in bold) would be largely interpreted as stressed because of the Spanish correlates of word-level (not intonational) on the syllable. A native English speaker, then, would have to perceive these correlates as stress to determine the meaning of the sentences.

Because of the frequency with which Spanish words differ only in stress, lexical access is affected strongly by stress in Spanish. Cutler and Pasveer (2006) describe Soto-Faraco, Sebastián-Gallés and Cutler's 2001 study on stress-based lexical priming, which found that in Spanish, the beginnings of spoken words with stress information prime words with the same stress pattern over words with a different stress pattern, even when all sounds are the same. English-speakers, on the other hand, do not benefit from this priming. Cooper, Cutler and Wales (2002) found that English speakers showed little change in processing time for priming words with different stress patterns.[2] This difference likely results from the differences in use of contrastive stress in the two languages. English, like Spanish, has contrastive stress, but minimal pairs are rarer and usually cross word categories. These minimal pairs "are predictable and almost always semantically related" (Saalfeld, 2012: 285). For example, ***per**mit* and *per**mit*** have similar definitions - a **per**mit may per**mit** someone to do something - but the first is a noun and the second is a verb. Even ***con**tent* and *con**tent***, which are not semantically related, differ in categories; the former is a noun, and the latter is an adjective. Because verbs and nouns are syntactically very different in English, it is less likely that L2 speakers would have to use only stress to differentiate an entire sentence. Cutler (1986) even found that stress-based minimal

[2] Cutler and Pasveer (2006) find English's vowel reduction responsible for this difference between English and Spanish because English contrasts the vowels in stressed and unstressed syllables.

pairs in English are processed as homophones (Saalfeld, 2012); that is, there is no need for English speakers to encode stress information to differentiate two words. Therefore a large difference exists between the two languages in the use of stress: while Spanish often uses stress to disambiguate words, English instead carries other syntactic differences in words that differ in stress.

That is not to say that there is no use for stress in English. Instead of differentiating between words, English employs stress to differentiate between one word and two. Word boundaries are often shown by stress. For example, ***green****house* and ***green house*** (in which both words are stressed, though prosodic stress may emphasize one word more), mean very different things, but can only be differentiated by the word boundaries formed by stress information. Here the presence of one stressed syllable per word informs the listener of word boundaries because the two-word phrase has stress on both words[3]. In contrast with English, Tyler & Cutler (2006) and as LaCross et al. (2016) both showed that Spanish did not strongly use stress in word boundary detection. Because Spanish has more use of contrastive stress than English, it is used more for lexical access than for word boundary detection.

Even if stress is recognizable enough to disambiguate words and aid in lexical access, stressing the wrong syllable can index the level of experience a speaker has with the language, or when the speaker started speaking it. Magen's 1998 study on pronunciation of L2 English as judged by native speakers showed artificially raising the pitch on stressed syllables in L2 speakers' utterances significantly affected how "native" the speaker sounds. When scoring the speech on a scale from 1 to 7, 1 being closest to a native speaker's speech and 7 being the farthest from a native speaker's speech, the score for lexical stress decreased from 4.51 to 4.29 after editing the pitch on the stressed syllables (p. 399). This is a significant change, showing that stress, when not pronounced as native speakers would expect, contributes to perception of "foreignness". Accents often affect a group's perception of a speaker's identity; a stronger accent, regardless of its actual reasons for existing, may give the impression of a lack of familiarity with the language (and, by extension, the culture), affecting societal treatment of a speaker. Accents are often confused with a lack of fluency in a language, even though they do

[3] See Taft (1984) and Cutler & Norris (1988) for experiments based on English stress-based word boundary perception.

not necessarily correlate with each other, so that a speaker with a stronger accent may not be able to get the same jobs or services as a speaker without one.

1.2 Correlates of Stress

There are three main correlates of stress production: duration, intensity, and pitch. Stress may also be expressed through phonological vowel reduction of unstressed syllables. A language may use any or all of these correlates. The duration correlate is usually expressed through lengthening the vowel in the stressed syllable, but some languages express stress by lengthening a syllable's consonant.[4] Pitch usually refers to a rising pitch on the stressed syllable, though technically any difference in pitch may mark stress. Note that "pitch accents" are considered in many articles to refer to prosodic stress, or in tonal or otherwise languages as another phonemic contrast in the language. Intensity usually refers to an increase in volume on the stressed syllable.

Spectral tilt is another method of examining intensity. Because increased effort in the glottis when intensifying a syllable causes it to open and close unevenly so that it takes a much shorter time to close than to open, higher frequencies are intensified more than lower frequencies. This "intensity distribution" (Sluijter & Van Heuven, 1996: 2472) is called "spectral tilt", and may be perceived as increased volume because of the "rising spectrum" of intensity:

> Overall intensity is not the only valid operationalization of increased physiological effort; we should at least consider intensity distribution, or spectral tilt, as well. Spectral tilt, in contradistinction to overall intensity, is not easily obscured by environmental factors, so that this operationalization of greater vocal effort seems communicatively more robust than overall intensity. (Sluijter & Van Heuven, 1996: 2472)

Regardless of the distinction between intensity and spectral tilt, however, they both require increased glottal pressure, as does pitch. Therefore, Lehiste (1970) explains that intensity and pitch are physiologically linked because pressure under the glottis, which creates higher intensity, also makes the vocal folds vibrate more quickly, causing higher pitch if the tension of the vocal folds is not changed.

English and Spanish differ in the correlates of stress they use. The linguistic community largely agrees on duration as the primary correlate of English. However, this consensus, which

[4] See Lunden et al. (2017) for a discussion of the correlate of consonant-based duration.

began in the 1990s, was a large change from the original thought, stemming from Fry's influential 1958 study (in addition to his 1955 one), which said that pitch was a primary correlate in English. However, Beckman and Edwards (1994) determined that one of Fry's methods of data collection, namely list elicitation, had caused each utterance to have its own phrasal contour. Therefore, prosodic stress had interfered with Fry's data. Beckman and Edwards concluded that pitch was secondary to duration in non-prosodic stress, a conclusion that remains today.

The conclusions on Spanish stress correlates, however, are not nearly as clear, despite efforts to isolate word-level stress from prosodic stress. Ortega-Llebaria and Prieto (2010) compared declarative sentences with reporting clauses[5], what Kim (2015) calls "parenthetical sentences" (p. 108), to eliminate the influence of prosodic stress correlates (Ortego-Llebaria & Prieto, 2010: 81). They found that Spanish had a minor but still present contrast in intensity and less of a contrast in duration between stressed and unstressed syllables than Catalan, though duration was still an important cue (p. 85, 88). In their study, however, they criticize Llisterri et al.'s 2003 article for only using declarative sentences, therefore not separating out prosodic stress. The 2003 article concluded that no correlate can cue stress on its own, and that pitch is necessary to cue stress unless both duration and intensity are present.

The most recent study on Spanish stress correlates was in Vogel et al.'s 2016 article, which concludes that pitch and pitch change relative to other syllables is the primary correlate of stress in Spanish in words not in focus (p. 138). In contrast with Llisteri et al. (2003)'s and Ortego-Llebaria and Prieto (2010)'s articles, which did claim at least a small influence from duration, Vogel et al. found that the stressed and unstressed vowels of words not in focus do not differ largely in duration; the only significant difference is between stressed vowels in focus and stressed vowels out of focus (p. 139). As the most recent exploration of Spanish word stress, as well as one of the only studies to fully separate prosodic and word-level stress, this article forms the basis of this thesis's hypotheses.

[5] Reporting clauses are clauses that "designate… the subject and [are] spoken with flat intonation, like 'John said' in John said, 'I'm going to the shops.'" (Ortega-Llebaria et al., 2013: 186). Declarative sentences are defined in this article as "refer[ring] to a broad focus utterance spoken with (non-contrastive) pitch accents on each stressed syllable, such as 'Mary is coming' when spoken as an answer to 'What's going on?'." (Ortega-Llebaria et al., 2013: 186).

The difference of vowel reduction in the two languages cannot be ignored, even in a study that attempts to avoid the distinction. English features phonological vowel reduction of unstressed vowels to schwa, a centralized vowel. Therefore, the first vowel in ***pro**duce* is full, but in *pro**duce*** it is a schwa. The issue of syllable weight, mentioned above, causes the second syllable of the former not to reduce completely, but in a word with all light syllables, every unstressed vowel reduces, such as in *ba**na**na*. Spanish does not have this reduction; all vowels remain phonologically the same in unstressed position. However, Nadeu (2014) points out that some phonetic vowel alteration in unstressed vowels does exist. This change likely occurs because stressed vowels usually require less articulatory effort (see Sleujter and Van Heuven, 1996 and Lehiste, 1970) and therefore some phonetic centralization occurs. Because there is no systematic or categorical reduction, however, the reduction cannot be called phonological.

Spanish may also differ from English in its differentiation between different stressed vowels. Ortega-Llebaria et. al found in their 2008 study that duration and intensity cues are more magnified on the stressed [a] than on [i], stating the theory that duration is only used for the [i] vowel if intensity cues are absent, likely because [i] vowels tend to be shorter than [a] ones: "Spanish speakers take from the signal whatever cue is available, which depends on the vowel type" (p. 4). The [e] vowel used in the present study balances this contrast by averaging the two vowel heights.

In addition to determining that Spanish and English stress production as well as perception appear to differ for native speakers regarding their own language, linguists have also compared speakers of both languages. Ortega-Llebaria et al. (2013) played reporting clauses for native speakers of the two languages, finding that Spanish speakers heard durational differences more easily than English speakers. When they played declarative sentences, however, both groups could perceive the stress. Ortega-Llebaria's 2010 article (described above) established that reporting clauses were at less risk than declarative sentences of prosodic influence, so the results from reporting clauses are likely more representative of word stress. When discussing these results, Ortega-Llebaria et al. (2013) mention Delattre (1966)'s finding that that the length of stressed vowels were six times that of unstressed vowels in English but only three times that of unstressed vowels in Spanish. They therefore predict that Spanish speakers' familiarity with smaller differences in stress "may be conducive to Spanish speakers being more successful that English speakers at perceiving these small duration differences in reporting clauses or post-focal contexts as potential cues to stress, suggesting a possible effect of cross-language duration differences in stress perception" (p. 189).

There is much less research about learning stress correlates than about learning stress systems of other languages[6], and even less based on Spanish and English comparisons[7]. Kim's 2015 study is one of the only ones that explores the perception of stress through correlates rather than through the position in a word. She focused on Spanish heritage speakers who are more dominant in their second language, English, exploring the influence of the heritage language on the dominant language. She found that heritage speakers of Spanish, even though they were more dominant in English than Spanish, still retained the ability to perceive stress in Spanish words more accurately than native English speakers. Unfortunately, the correlate used in the Spanish stressed syllables was duration, which is not currently considered the primary correlate of Spanish stress. However, given the previous research on Spanish speakers' stronger ability to perceive durational cues, this study may be evidence for the influence of hearing a language as a child, even a non-dominant one, on stress correlate perception in adulthood, regardless of the primary correlate of production. Given that babies exposed to languages without contrastive stress cannot distinguish stress by nine months old (Mattock & Burnham, 2006), one could assume that more nuanced abilities to perceive stress correlates may also be influenced at a young age.

1.3 Present Experiment

When developing the current experiment, the stimuli were chosen based on previous research. First, the importance of a lack of context for the words was established. Eriksson et al. (2002) discovered the influence of contextual expectations of speakers when perceiving stress. English speakers who did not speak Swedish, along with Swedish speakers, were given stimuli that included words with stress in positions that contradicted where they would have typically occurred in the language. Fluent Swedish speakers were not as able to perceive this stress as non-speakers of Swedish because they predicted stress based on context. In order to avoid context-

[6] See Wang (2008) and Yu & Andruski (2010) for information on stress perception for Chinese-speaking English L2 learners.

[7] See Archibald (1993) for information on transferring stress patterns from Spanish to English.

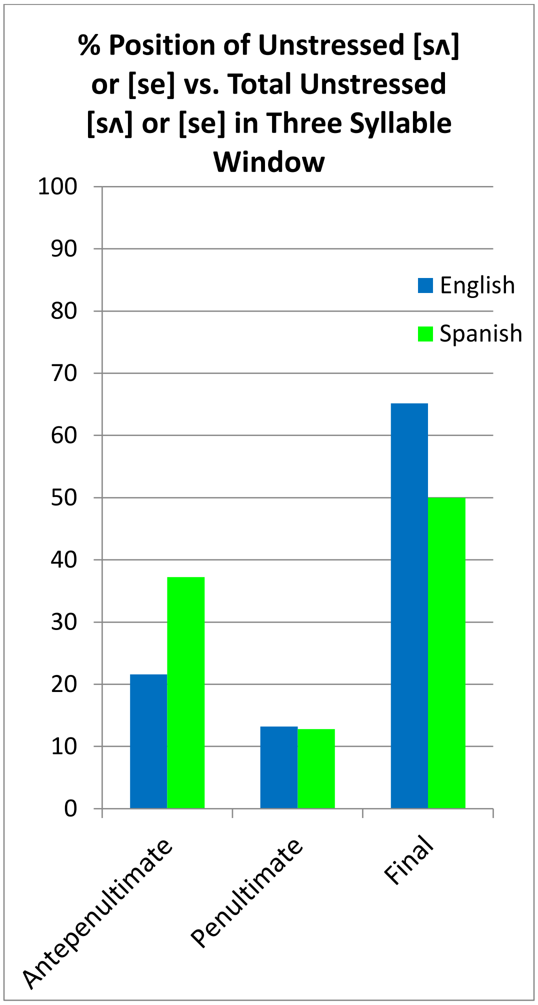
based stress predictions in the current study, the tokens used were nonce words. Lack of context or real words does remove external validity, meaning the study does not exactly show how speakers perceive real speech in their everyday environments; however, the present focus on perception of stress correlates would be impeded with this outside information. The possibility of analogy was also eliminated because it was previously shown to aid in stress perception. Bullock and Lord (2003) found that L2 learners of Spanish often use analogy to determine where stress would be in a word, and when the L2 vocabulary is not sufficient to find such an analogy, they use their L1 lexicon. Even though these findings were based on production of written words, some analogy could be used in a perception study. The current study eliminates the use of analogy by only using one nonce word form that is not similar to any real words. Lastly, the light syllables eliminate the risk that L2 English speakers might associate some heavy syllables with stress, because Peng and Ann (2001) found that diphthongs often attract stress in L2 English speakers' speech. The syllables are all the same so vowel quality will not influence the determination of stress location, and the vowel is not one shown in the aforementioned study to attract stress in production.

In order to test the validity of the experiment, the frequency of the syllables used in English and Spanish was examined. Vitevitch et al. (1997) showed that "[t]here [is] no interaction between phonotactic probability and stress, suggesting that participants treat… these two sources of information separately in making their judgments" of the "goodness" of phonotactically legal nonsense words (p. 60). However, judgments about acceptability differ from trying to find the location of stress. In the case of this study, the initial consonant of a syllable may still have some effect on stress perception.

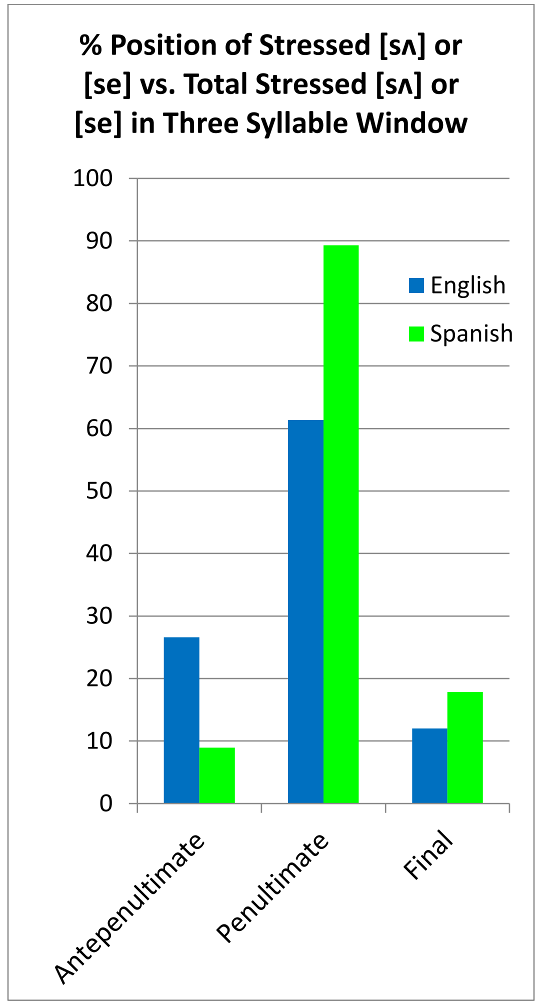
Section 2.2 describes the stimuli for the current study. The frequency of the syllable of the stimuli was examined to ensure that the two syllables were comparable. Two corpora were used to examine these frequencies. For Spanish, BuscaPalabras was employed, and for English, the Carnegie Melon University Pronouncing Dictionary was used; both sources were tagged for stress. This analysis found that the syllables [se] in Spanish and [sə] in English occur with similar levels of frequencies in different positions in the three-syllable window:

Figure (1): Frequency Comparisons of Stimuli Syllable

(a)



(b)



When the syllable is present and stressed, it is most likely in both languages to be in the penultimate position rather than antepenultimate or final. When the syllable is present and unstressed, it is most likely in both languages to be the final syllable. The only difference between the order of most frequent to least is that in Spanish, the stressed [se] is slightly more likely to be final than antepenultimate, while the stressed English [sə] is slightly more likely to be antepenultimate than final. The syllable [sə] does occur less frequently as a stressed syllable than [se] does in Spanish, but there are more vowels in English than Spanish, so even if a typically-reduced vowel had not been chosen, the frequency of the syllable when compared to total stressed syllables would likely have been lower than for Spanish. The English vowel also occurs significantly more often when compared to other unstressed vowels than Spanish [se], but the presence of vowel reduction in English allowed this vowel to be the only one that would not change in completely unstressed syllables, therefore also causing this frequency difference.

1.4 Research Questions

When a speaker can speak two languages, Flege's Speech Learning Model (2003) states that there is one "phonetic space" for both languages (Kim, 2015: 107). If this "space" applies to suprasegmentals as well, then the stress correlates of the two languages would be likely to be used overlappingly. In this case, perception would likely mirror production, and bilingual speakers would have access to both stress correlates. This hypothesis, however, does not tackle the issue of variation in how long speakers have been learning or speaking the other language. Would native speakers of both languages be able to use both languages' correlates as someone who just learned a second language? Even if there is an overlapping space, can bilingual speakers code-switch between correlates of stress, just as they can code-switch phonology?

Four research questions were developed:

1. How do bilingual speakers with differing dominant languages differ in their perception skills?
2. What is the threshold of changed features required for a native speaker of either language to identify stress?
3. How do bilingual speakers differ in perception depending on the language they're hearing or expecting to hear?
4. How does age of acquisition (in relation to the first critical period[8]) affect language-specific stress perception?

This research aims to answer these questions in order to better understand the acquisition of non-native languages. Three hypotheses align with these research questions:

1. Perception will match production correlates of stress for particular languages.
2. Bilingual speakers will perceive stress with the correlates of their dominant language, rather than the correlates of their non-dominant language.

---

[8] See Gleitman and Newport, 1995.

3. Bilingual speakers will change the correlates for which they are listening depending on the language they are expecting to hear.

The following experiments compare the results of bilingual Spanish and English speakers who are native in one of the languages. They hypothesize a connection between perception and production; therefore, the studies that show that Spanish speakers use pitch and English speakers use duration and intensity on stressed syllables would imply that perception of these correlates would differ depending on the native language spoken. The results from the future study can be used to look at bilinguals' changes in perception depending on which language they are expecting to hear, as well as, given that one experiment will target the native language, to look at the difference between English and Spanish speakers in general (by looking only at the native language portion of the experiment).

2. Methods

2.1 Participants

The participants were 14 college students, ages 18 through 21 from The College of William and Mary (W&M) in Williamsburg, Virginia (average age = 19, M=3) and 14 college students, aged 18 through 25 from La Universidad Peruana de Ciencias Aplicadas (UPC) in Lima, Peru (average age 20, M=5)[9]. Most, but not all, students from William and Mary received participation credit for an introductory Linguistics or Psychology class; none of the students at UPC received credit.

Because the students in the United States were exposed to English as the dominant language of their environment, and the students in Peru were in contrast exposed to Spanish, the two groups were analyzed as opposites, with the students in Peru asked more about their exposure to English, and students in the United States asked about their exposure to Spanish. When comparing the groups, "non-dominant language" refers to English for students at UPC and Spanish for students at W&M. The "target language" was considered to be the non-dominant language of the country in which the students attended college.

[9] Four students at William and Mary and one student at UPC were left-handed.

All of the students at both universities stated that they learned their country's dominant language before the first critical period[10]. Nine students at William and Mary identified as native speakers of their country's non-dominant language (Spanish)[11]. In contrast, only one student at UPC identified as such. Four William and Mary students and six UPC students learned their country's non-dominant language between seven and twelve years old. One William and Mary student and four UPC students learned the language between thirteen and eighteen years old. The rest[12] -- nine William and Mary students and three UPC students -- learned the language at age six years older or younger.

Nine William and Mary students learned Spanish at home, whereas only three UPC students learned English at home. None of the students at UPC learned English with complete immersion, though ten learned it in the target language, while ten students at William and Mary learned English through complete immersion, and only one was taught in the target language[13]. Six students in Peru learned English at a separate school from their usual studies, where they went specifically to learn English; no William and Mary students learned Spanish in this way. In general, the students in Peru appeared to have learned the non-dominant language in a more scholarly setting than the students at William and Mary, who largely learned it as heritage speakers.

Despite the disparities in native languages - the majority of the students at William and Mary were native speakers of both languages, while the overwhelming majority of the students

---

[10] All 14 students in Peru identified as "native" Spanish-speakers, while only 12 out of 14 students at William and Mary identified as "native" English-speakers; however, the two students who did not self-identify as "native" both stated that they learned English before age six (which is consistent with what is commonly taken to be the end of the first critical period; see Gleitman and Newport, 1995).

[11] One participant also stated that they spoke Korean natively.

[12] One student at UPC did not answer the age at which they started learning English, but then stated that they learned it in primary school.

[13] Three students at William and Mary and four students at UPC were taught in their country's dominant language.

at UPC were native speakers of Spanish only - the self-identified skill levels in the country's non-predominant languages were extremely similar:

Table (1): Average Scores from 1 to 7 for Country's Non-Dominant Language

	Reading Comprehension	Oral Competency	Listening Comprehension
USA	5.64	6.14	6.07
Peru	5.93	5.43	5.71

Table (2): Standard Deviations for Country's Non-Dominant Language

	Reading Comprehension	Oral Competency	Listening Comprehension
USA	1.08	0.949	0.829
Peru	0.475	1.09	0.825

The fairly low standard deviations show that among each school's competency the students gave themselves fairly similar scores, and the extremely close average scores show that the populations were fairly similar in skill level and true bilingualism. The highest average score at UPC was for reading, while reading was by far the lowest score at William and Mary. This difference is logical given that 10 out of 14 William and Mary students learned Spanish at home, but only three UPC students learned English in this setting. Learning a language in a school setting would likely focus more on reading comprehension than would learning the language in a non-school setting.

2.2 Stimuli

Four-syllable strings, used in two different studies, were constructed by isolating syllables from a recording of a female bilingual speaker, age 21. The recording consisted of question and answer pairs in Spanish and English that contained words not in focus that began with two stressless target syllables, [sə] for English, and [se] for Spanish. These target syllables were repeated to create four-syllable strings of repeated syllables for each language. Either the second (antepenultimate) or third (penultimate) were manipulated in Praat (Boersma, Paul & Weenink, 2016) to produce one stressed syllable per string. This stressed syllable was altered in duration, pitch, or both. There were seven possible values for the stressed correlate and two possible locations for the stress, creating 42 levels for each language. Duration levels' "stressed" syllables increased by one wavelength after the initial level, which was two wavelengths above the baseline (14 wavelengths, from the original recording with pitch slightly smoothed). Pitch levels' "stressed" syllables increased by 5 Hz on the pitch peak, starting 5 Hz above the baseline (210 Hz)[14]. The "unstressed" syllables contained the baseline values for pitch and duration, and when a correlate was changed for pitch, the other correlate remained at this baseline (except for the combination levels). The combination levels' "stressed" syllables used the same values for pitch and duration but included both for each level (e.g., starting with level 1 duration combined with level 1 pitch). Each "stressed" syllable was present four times in the study for each language: twice in the penultimate position, and twice in the antepenultimate position. The following table shows the values for pitch and duration of the stressed syllables:

Table (3): Levels of the Independent Variables

Level	Duration of Vowel (wavelengths)	Duration of Vowel (ms, English/Spanish)	Pitch (Hz)	Combination (wavelengths-Hz)
Baseline	14	67.3/69.2	210	N/A
1	16	76.9/78.8	215	16-215
2	17	81.7/83.6	220	17-220
3	18	86.5/88.4	225	18-225
4	19	91.3/93.3	230	19-230
5	20	96.1/98.1	235	20-235

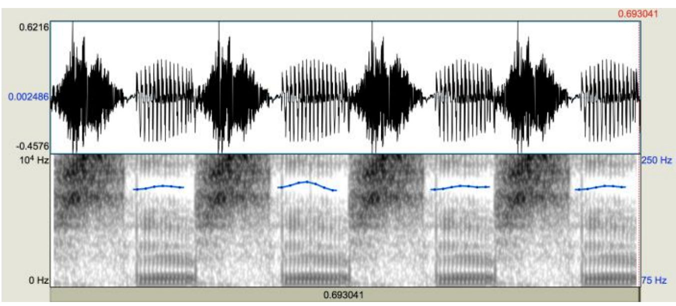
[14] Pilot studies were conducted in spring and summer 2017 to test the correlate ranges. It was found that the highest levels had unnatural buzzing and the correctness and goodness scores dropped off for the top few levels, so the range was adjusted accordingly.

6	21	101.0/102.8	240	21-240
7	22	105.7/107.6	245	22-245

The baseline was based on the English syllable, so the Spanish syllable was changed in pitch slightly to match. The intensity of each syllable was 72.7 dB for English and 75.8 dB for Spanish. Because of the original length of the two vowels, all of the English durations were less than 2ms under the Spanish durations for the same level.

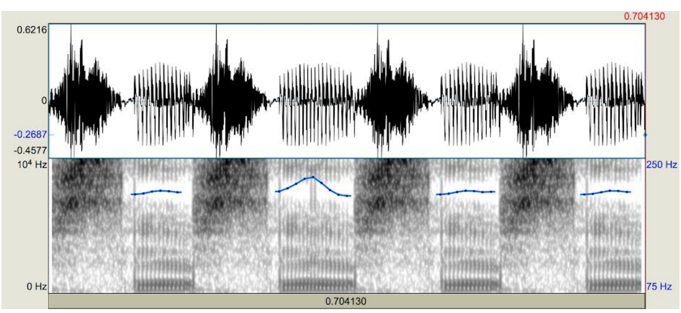
The following spectrograms highlight the differences between the levels, and show the pitch and duration qualities of the combination correlate. The strings are from the Spanish section, and all have antepenultimate stress.

Figure (2): Spectrogram of Level 1



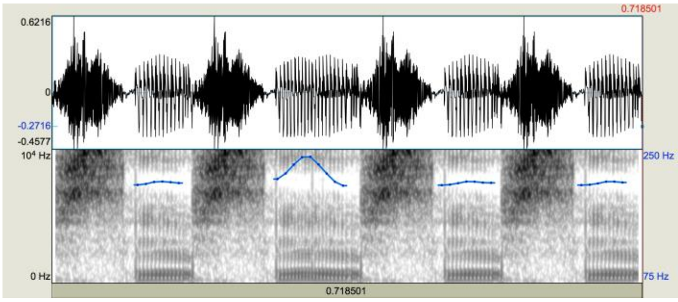
Antepenultimate position, Combination correlate

Figure (3): Spectrogram of Level 4



Antepenultimate position, Combination correlate

Figure (4): Spectrogram of Level 7



Antepenultimate position, Combination correlate

Because these examples were ones in which the stressed correlate was changed for both pitch and duration, the bell-shaped line gets to the tallest point in Level 7, and the syllable's vowel is longest in this level, as well.

2.3 Procedure

First, recorded descriptions of stress were played for participants in their country's dominant language. The Spanish descriptions were recorded by a male bilingual professor from the College of William and Mary who studied in Lima, Peru, and the same description translated into English was recorded by the researcher. Students were provided two examples of stress per language: **lla**me and lla**mé** (with stress on the penultimate and final syllables, respectively) for Spanish and **per**mit and per**mit** (with stress on the penultimate and final syllables respectively) for English. Students were asked in the recording for confirmation of their understanding, and no students asked for clarification during this section of the instructions. Then, participants read a brief passage on Qualtrics and answered three simple reading comprehension questions (see appendix D). Every alternate student read a passage and answered questions first in Spanish and the other half of students read a passage and answered questions first in English, regardless of their country's dominant language.

When the participants finished the questions, the researcher played another set of more specific directions in the same language of the reading comprehension questions, and more examples per language were given: a**pós**trofe and 'mari**po**sa' (with stress on the antepenultimate and penultimate syllables, respectively) for Spanish and a**spa**ragus and 'Massa**chu**setts' for English. During these recordings, students were played examples of penultimate and antepenultimate stressed strings using the syllables [ba] and [ga]. After the recording ended, the students saw a summary of the instructions on the screen in front of them. They then began the study in the language of the previous instructions. During the study, students used their middle and index fingers to identify the location of the stress in the syllable strings they heard by hitting "2" or "3" with their left hands if the stress was on the antepenultimate or penultimate syllable, respectively. After pressing the key, they then clicked a number on a goodness scale at the bottom of the screen with their right hand; 1 marked least clarity and 5 marked most. Reaction times between the end of the recording and the selection of the stress location were collected using Praat.

When finished, participants read a passage in the other language and again answered reading comprehension questions. They repeated the process of recorded instructions followed by the study. Finally, they filled out an online language questionnaire in the dominant language of their university's country regarding their background learning their second language and their abilities in the language.

Beyond the 28 participants discussed, data was excluded from participants with correct stress identifications below the threshold of 100/168, because a minimum of 100 correct corresponds to a one-tail binomial calculation nearest the threshold of p=.01 (exact value .0083). One participant only finished half of the experiment and was therefore excluded, as well. Seven participants' data was excluded from UPC and eight participants' data was excluded from W&M.

3. Results

SPSS was used to fit three Generalized Linear Models with three dependent variables – RESPONSE, REACTIONTIME, and GOODNESS rating – and five independent variables – *language* (the language of instruction, corresponding to the syllable in the string; two levels: English and Spanish), *group* (two levels: UPC and W&M), *correlate* (three levels: duration, pitch, and combination), *level* (seven levels: 1-7), and *position* (two levels: antepenultimate and penultimate), as well as their interaction terms. *Participant*, nested in *group*, was treated as a blocking factor. While *language* was included in the RESPONSE and GOODNESS models so as to answer an initial experiment hypothesis and look into UPC participants' feedback, it was excluded from the REACTIONTIME model because it was not meaningfully significant in either of the former models. Appendix A contains the significances of each independent variable in tables for each dependent variable.

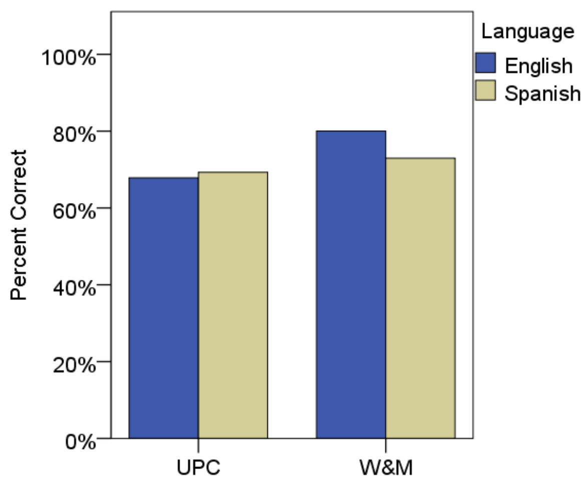
Data for REACTIONTIME was cut if it was above 5ms and under .5ms, but was kept in for the other dependent variables. REACTIONTIME and GOODNESS were only analyzed for the correct RESPONSES. Post-hoc pairwise comparisons, using Fisher's least significant difference (LSD) adjustment, were run for all interaction terms, and p-values reported below derive from these tests' significance *level*s ($p \leq .05$ marks significance).

3.1 Influence of Language *on* RESPONSE

The factor *language* was not significant in the overall linear model for the dependent variable RESPONSE (p=.405), nor was the interaction of *group***language* (p=.455) (although it had a significant reaction with *position* (p=.022)).

However, comments after the study from UPC participants pointed to a higher difficulty when the *language* was from the *group*'s non-dominant one[15]. Therefore, this facet of the study was still analyzed. The graph below shows each group's scores on each section. The outside bars represent the theoretically more "difficult" sections.

Figure (5): *Group*'s RESPONSE for *Language*



Despite the UPC participants' comments, there was no significant difference in score for *language* within each *group* ($p \geq .259$). However, within each *group*'s *language*s, broken up by

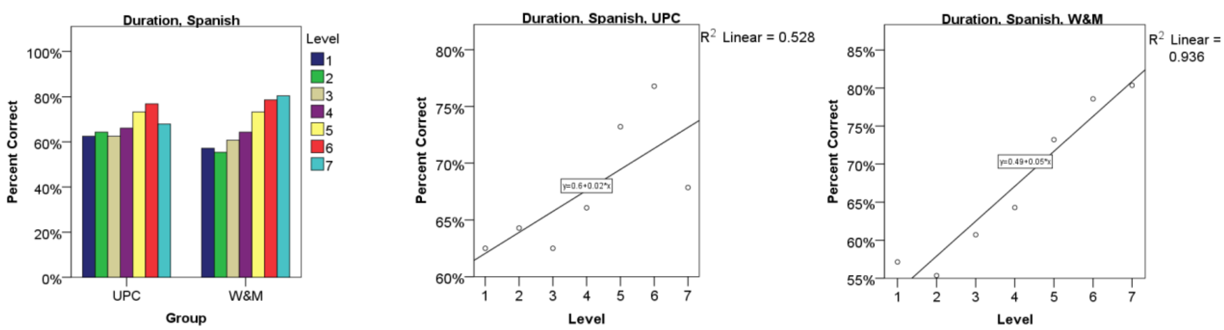
[15] Again, while other languages are clearly present in both countries, but the language most commonly used at each university was considered that *group*'s 'dominant language'.

correlate, the perceived difficulty may have occurred only over certain *level*s; *correlate***level* (though without *language*) did have borderline significance (*p*=.064). Therefore, the *level*s of each *group*'s *correlate* for each *language* were broken out. The percent correct values for each *level* were placed in a scatterplot, and a line of best fit was calculated for each one. This line showed the general trajectory of improvement over the *level*s. At low *level*s, perception of frequency (Hz) is linear, and the pilot studies conducted also showed a linear increase for 'duration' and the 'combination' *correlate*s; therefore, a line was appropriate for all of the *correlate*s' increases.

The following two scatterplots (Figures (8) and (9)) reflect the data for *correlate* 'duration' in *language* 'Spanish' (Figure (7)), divided by *group*. We see that there is an upwards trend for W&M, corresponding to more correct RESPONSES as the duration of the stressed syllable increased:

Figure (6): Lines of Best Fit and R^2 Values for *Language*, *Correlate*, and *Group*

(a) (b) (c)



The slope of the best fit for W&M, shown on the right of the bar graph and in the right scatterplot, is notably steeper than that of UPC, shown on the left. While this slope can be estimated from the bar graph, the best fit line provides a numerical verification. In addition, the R^2 value in the scatterplots show how well the *level*s adhere to a linear increase. Therefore, the less consistent scores from UPC are shown to have a lower R^2 value for its line of best fit.

A table of slope values for the lines of best fit, along with their R-values, were placed in Table 1 below. The following table shows the slopes and R^2 values for each *group* and *correlate*, separated into the two *language*s of instruction.

Tables (4) and (5): *Language* Slopes of Best-Fit Lines and R^2 Values for RESPONSE

(4) ‘English’

	Duration		Pitch		Combination	
	Slope	R Value	Slope	R Value	Slope	R Value
UPC	0.02	0.445	0.01	0.248	0.05	0.641
W&M	0.04	0.837	0.04	0.614	0.05	0.782

(5) ‘Spanish’

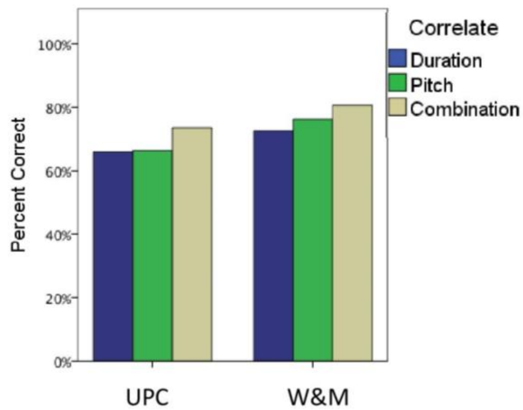
	Duration		Pitch		Combination	
	Slope	R Value	Slope	R Value	Slope	R Value
UPC	0.02	0.528	0.04	0.78	0.05	0.887
W&M	0.05	0.936	0.02	0.701	0.02	0.413

The ‘duration’ *correlate* in each *language* featured a notably steeper and more linear improvement for W&M than for UPC. The results for ‘pitch’ differed based on *language*. For the ‘English’ section, W&M had a steeper slope and more linear movement than UPC, but for the ‘Spanish’ section, UPC had a much steeper incline for improvement. ‘Combination’ in ‘English’ showed similar results for each *group* with a notably higher slope than the other *correlate*s within the same *group*, as well as a fairly linear progression. Although the Generalized Linear Model Regression did not find significant *language*-based effects, the R^2 values in each *group*’s dominant *language* do represent a more linear path than the opposite *group*’s. We see this phenomenon with W&M having higher R^2 values than UPC in ‘English’ and UPC having higher R^2 values than W&M, though only for ‘pitch’ and ‘combination’.

3.2 Perception Differences in RESPONSE

Factoring out *language* (because it was not significant) and combining the *correlate*s' *level*s allows us to compare the results for each *correlate* divided by *group*. The overall effect of *correlate* was significant (*p*=.003), while its interaction with *group* was not (*p*=.271). The following graph displays *correlate*s broken out by *group* despite this lack of significance because separating the populations was the premise of this study.

Figure (7): *Group*s' RESPONSE for *Correlate*



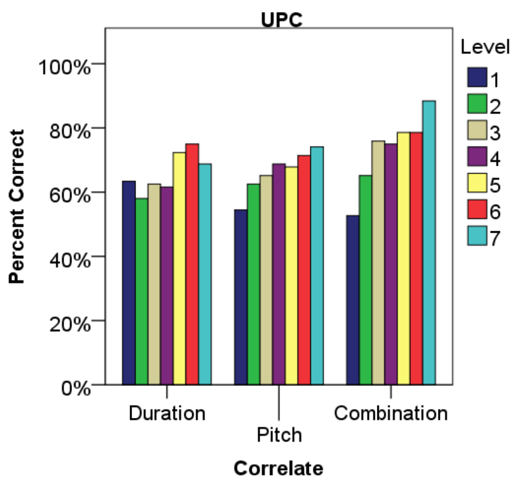
UPC was significantly better at identifying stress when indicated by 'combination' compared to 'duration' (*p*=.003), while W&M was not (*p*=.081); UPC was also better at identifying stress when indicated by 'combination' compared to 'pitch' (*p*=.004), while W&M was not (*p*=.548). There were no significant differences between 'pitch' and 'duration' for either *group* (*p*=.952 for UPC, p=.307 for W&M).[16]

While the interaction of *group***correlate***level* was not significant (*p*=.113), we do see some differences in the linearity of the increase of *RESPONSE* as the *level* of *correlate* is increased for the two *group*s. Therefore, the *level*s in each *correlate* were then displayed for each *group*, showing (as expected) a general increase as the *level*s increased. The averages shown in the graph above are broken down by *level* in the graphs below; each cluster of bars represents one *correlate* for each *group* and is in the same order as the graph above: 'duration', 'pitch', and 'combination' for UPC on the left and the same *correlate*s for W&M on the right.

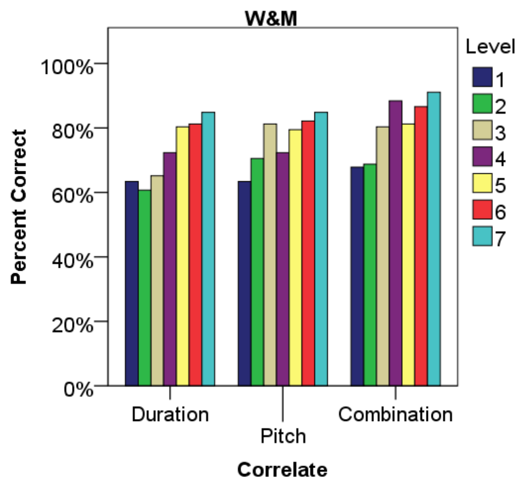
[16] Overall, the W&M participants performed somewhat, though not significantly (*p*=.071) better than the UPC students, getting 76.5% correct versus UPC's 68.6% correct.

Figure (8): *Groups*' RESPONSE for *Correlates' Levels*

(a)

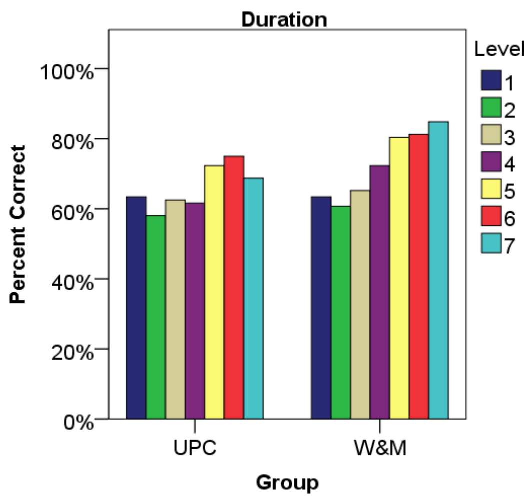


(b)



When these results are broken out by *level*, they show a notable distinction in behavior for 'duration' especially: there was notable linear improvement for W&M, whereas UPC did not show the same linearity, featuring relatively flat data. The bar graph below pulls out this data for 'duration' to compare side-by-side. We can clearly see the shallower slope for UPC and more linear trajectory for W&M:

Figure (9): *Groups*' RESPONSE for *Levels* of 'Duration'



These observations are reflected in the slopes and R^2 values for all of the countries' *correlate*s: with a slope of .04 for W&M and .02 for UPC in 'duration' and an R^2 value of .917 for W&M and .555 for UPC, the W&M data is steeper and more linear than that of UPC. The rest of these values are presented below.

Table 6: Slopes of Best-Fit Lines and R^2 Values for RESPONSE

	Duration		Pitch		Combination	
	Slope	R^2 Value	Slope	R^2 Value	Slope	R^2 Value
UPC	0.02	0.555	0.03	0.897	0.05	0.85
W&M	0.04	0.917	0.03	0.735	0.04	0.791

'Duration' is the only *correlate* with such a notable difference, in which W&M clearly improved more than UPC.

3.2.1 Signal Detection Theory

Another method of quantifying this data is using Signal Detection Theory. Signal Detection Theory is a psychological theory in which experimenters "manipulate the presumed decision criterion through… aspects of experimental procedure in order to expose the sensitivity factors that remain unchanged" (MacMillan, 2002: 44). In order to determine how much noise causes listeners to not hear a signal, for example, the experimenter varies the amount of noise and asks participants if they can hear a signal or not. For this study, the sensitivity factors (amounts of the correlates) are changed in order to find the decision criteria (which correlates are important for the UPC and W&M), effectively reversing the process. Signal detection theory divides responses into four categories, hits (when a participant correctly hears identifies the signal when it is present), misses (when a participant does not hear a signal when it is present), false positives (when a participant hears a signal when it is not present), and correct rejections (when a participant correctly does not hear a signal when it is not present).

The present study's results were analyzed using these terms as if 'antepenultimate' was the "signal".[17] The results were then analyzed to examine the sensitivity to the differences in position for the *group*s. The hits were defined as selecting antepenultimate when the position was antepenultimate, and false positives were defined as selecting antepenultimate when the position was penultimate. Then, the following equation was used to find the d' scores for the groups' results, where M_s is the mean of the signal distribution, M_N is the mean of the noise distribution, and z(H) is the z score of the proportion of hits to the total chances and z(F) is the proportion of false positives to the total chances (Macmillan, 2002):

$$d' = M_S - M_N = z(H) - z(F)$$

The z-scores of these values were calculated and multiplied by the rate of the hits and false positives. They were then subtracted to find the d' scores. When d' scores are higher, there is a larger sensitivity to the signal. The following table displays each group's scores for each correlate. The overall d' score for UPC was .9842, and for W&M was 1.4579. Therefore, W&M showed a higher sensitivity to the stress differences than UPC.

Table 7: d' Scores for *Correlate*s

	Duration	Pitch	Combination
UPC	0.845	0.866	1.2573

[17] See Ji-Young Kim (2015) for a similar use of signal detection theory to examine sensitivities to stress.

W&M	1.2218	1.4415	1.7349

All of the correlates were higher for W&M than for UPC, showing a higher sensitivity to the differences between the positions. Both countries increased in sensitivity from ‘duration’ to ‘pitch’ to ‘combination’. Therefore, both countries benefited from the presence of both correlates, rather than just one.

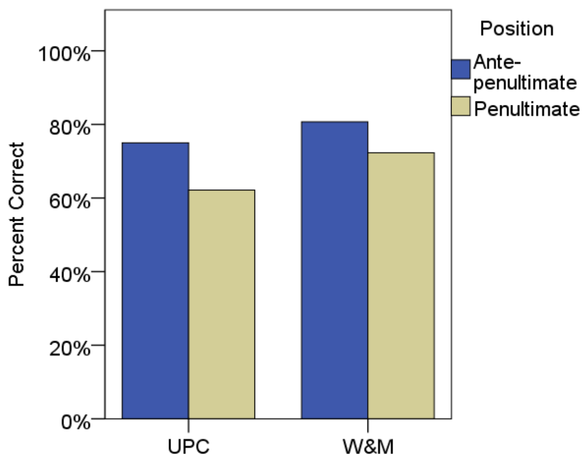
3.3 Comparing Correlates

The factor *correlate*, without interactions with *group*, was significant (*p*=.003). The absolute difference between each *correlate* should not be explored too deeply because the values of each *correlate* was fairly arbitrary, based on pilot studies that showed the most effective range of values. However, given that ‘combination’ consisted of both duration and pitch *correlate*s, this *correlate* can be compared to its separate components. ‘Combination’ had significantly better scores than for ‘duration’ (*p*=.001) and ‘pitch’ (*p*=.014), so even though ‘pitch’ cannot be compared with ‘duration’, we clearly see that the presence of more *correlate*s is significantly more helpful.

3.4 The Role of Position

The overall effect of *position* was significant (*p*<.001) as well as its interaction terms with *group* (*p*<.001), *language* (*p*=.022), *correlate* (*p*<.001), and *level* (*p*<.001) (as well as more complex interaction terms; see appendix A.) Both *group*s identified stress correctly more often when it was in the ‘antepenultimate’ *position* than in the ‘penultimate’ *position* (*p*<.001 for W&M and UPC). The following graph shows the differences in RESPONSE when *position* was ‘antepenultimate’ and ‘penultimate’. Using the signal detection theory labels above, the dark bars represent the hits, and the light bars represent the correct rejections.

Figure (10): *Groups’* RESPONSE by *Position*

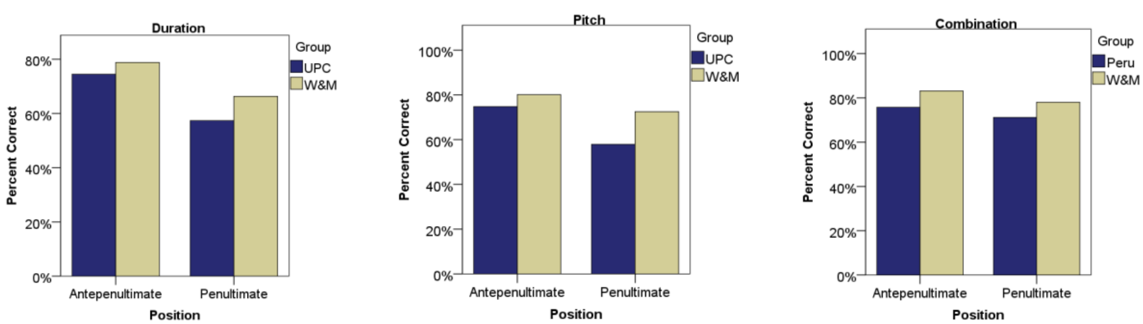


Each *group*'s scores for *position* when compared to the other *group*'s scores for *position* (e.g. comparison of dark gray bars in Figure (10)) was significantly different (p=.001 for 'antepenultimate', p<.001 for 'penultimate').

The interaction *group*position*correlate* was not significant (p=.493). However, because *position*correlate* was significant and dividing the populations was key to the purpose of this study, the following graphs remain divided by *group*, showing *position* for each *correlate*. This time the dark bars represent UPC, and the light bars represent W&M, to better compare the groups' scores in each position.

Figure (11): *Groups'* RESPONSE by *Position* for *Correlate*

(a) (b) (c)



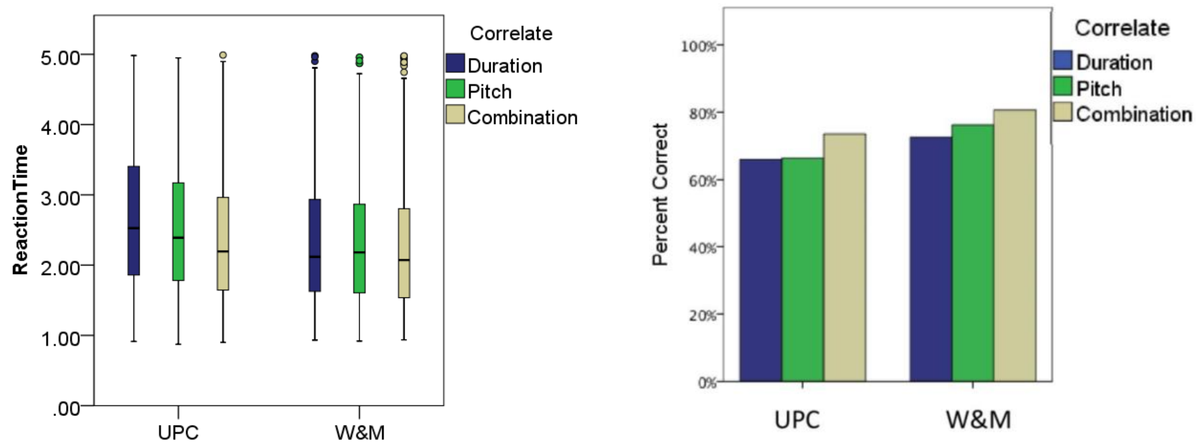
'Antepenultimate' and 'penultimate' within *correlate*, even when broken out by *group*, were all significantly different (p<.001 for each pairwise comparison). Therefore, no matter how the data is broken out, each *group* was better able to distinguish stress when it was in ‘antepenultimate’ *position*.

3.5 Results for ReactionTime

ReactionTime was the only implicit method of the study. As stated previously, ReactionTime data was trimmed, removing data under .5ms because of necessary processing time, and data above 5ms because it likely did not reflect the listener’s first instinct. Only the correct answers' ReactionTimes were analyzed. ReactionTime was significantly different for *group* (p<.001), *correlate* (p<.001), and *position* (p=.001). The results largely reflect those of Response, with slower ReactionTimes occurring when participants were less likely to identify the stressed syllable (lower Response). For example, the following graph displays ReactionTime for each *correlate* within *group*, next to a graph of the parallel Response data:

Figure (12): Comparing ReactionTime to Response for *Correlates*

(a) ReactionTime (b) Response

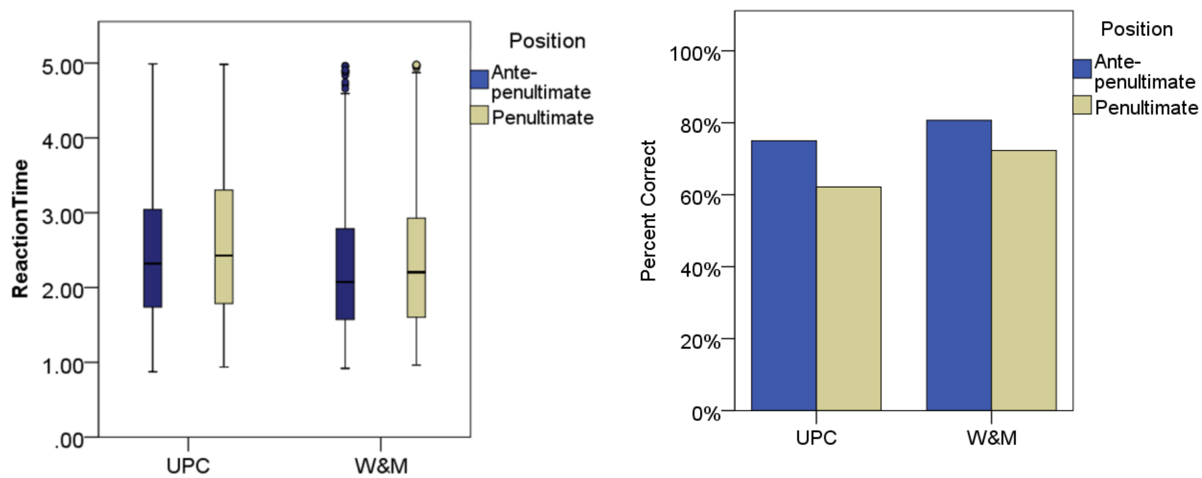


The increase in correct Responses from the ‘pitch’ and ‘duration’ to ‘combination’ is reflected in the lower ReactionTime needed for the correct Responses. Though the step-wise motion is not exactly the same, similar significant differences exist: for UPC between ‘combination’ and ‘pitch’ (p=.003 for ReactionTime) and ‘combination’ and ‘duration’ (p<.001 for ReactionTime). Unlike for Response, UPC’s ‘pitch’ and ‘duration’ were significantly different (p=.014), and W&M’s ‘combination’ and ‘duration’ were significantly different (p=.033).

Likewise, *position* data was parallel for REACTIONTIME and RESPONSE. There was no significant difference in the interaction of *group***position* (*p*=.758), but *group* remains broken out of the data to be consistent with the RESPONSE graphs in section 3.4. The following graphs show *position*, within *group*, for REACTIONTIME and RESPONSE side-by-side.

Figure (13): Comparing REACTIONTIME to RESPONSE for *Position*

(a) REACTIONTIME (b) RESPONSE

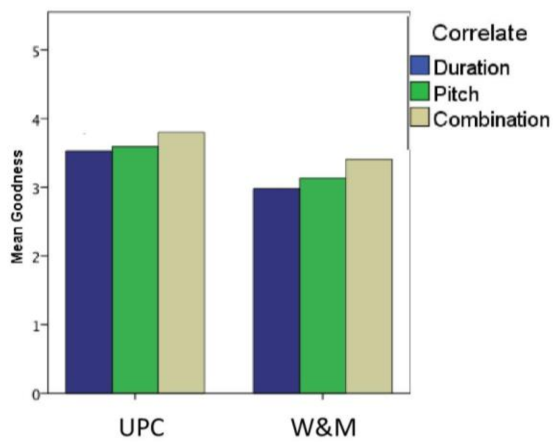


Both *group*s took a longer time to determine the location when stress was ‘penultimate’ (*p*=.012 for UPC REACTIONTIME, p=.023 for W&M REACTIONTIME). Hence, REACTIONTIME and RESPONSE match in which *correlates* require more processing effort.

3.6 Results for GOODNESS

There were two types of behavioral methods in the study; the GOODNESS rating was the second after determining the location of stress. After identifying the stressed syllable’s *position*, participants rated the clarity of the stress on a Likert Scale. As stated previously, only GOODNESS ratings for correct responses were collected. Unlike for RESPONSE and REACTIONTIME, both *group* and *language* were significant for GOODNESS (*p*<.001 for both). UPC's GOODNESS ratings were on average higher than the W&M's. *Correlate* was significant for GOODNESS (*p*<.001), as well. Once again, the interaction of *group***correlate* was not (*p*=.175), but *group* remains broken out because of the main research questions of the experiment. Note that the GOODNESS graphs have a maximum on the x-axis of 5 (though the scale was 1 to 7) in order to zoom in on the relevant data.

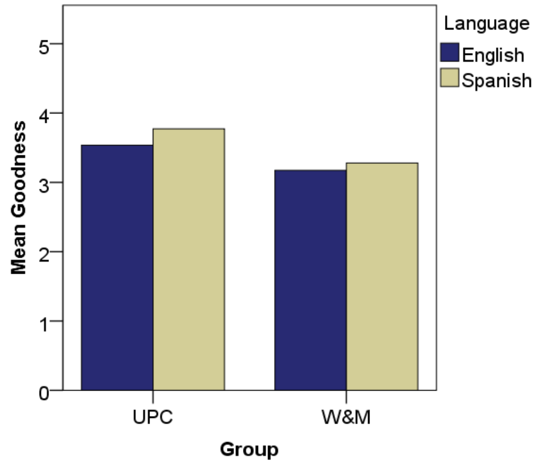
Figure (14): *Groups'* GOODNESS for *Correlates*



For UPC, differences between 'pitch' and 'combination' (*p*<.001) and 'duration' and 'combination' (*p*<.001), but not 'pitch' and 'duration' (*p*=.422), were significant. For W&M, differences between 'duration' and 'combination' (*p*<.001) and 'pitch' and 'combination' (*p*<.001), were significant, as well as 'pitch' and 'duration' (*p*=.025), differing from UPC. Participants found 'combination' to be a much more obvious indicator of stress.

Qualitatively, many UPC participants commented that 'English' was much harder to perceive. Though *language***group* was not significant for GOODNESS (*p*=.093), the graph below allows us to examine the participants' claims:

Figure (15): *Groups'* GOODNESS for *Language*



All of the pairwise comparisons were, in fact, significantly different. Within each *group*, there were significant differences between the *language*s (p<.001). 'Spanish', the dominant language of UPC, was significantly different from the same *language* between *group*s (i.e., 'Spanish' for W&M) (p<.001). 'English', the dominant language of W&M, was also significantly different from the same *language* between *group*s (p<.001). Instead of GOODNESS ratings favoring dominant languages, both *group*s favored 'English' over 'Spanish'.

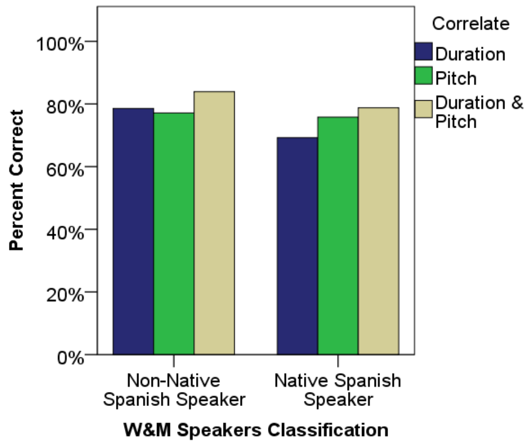
3.7 Spanish Native Speakers

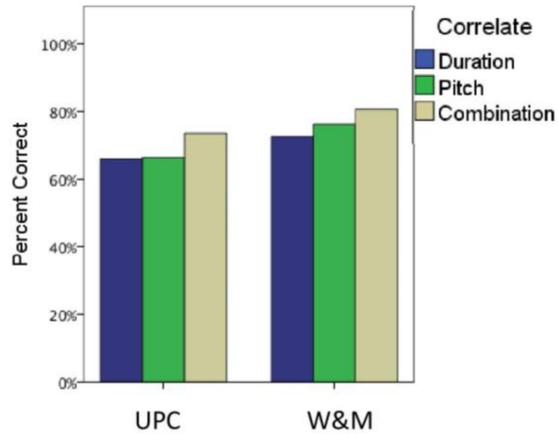
To address research question 4, examining the effects of learning a second language later than the first critical period, the W&M speakers who were natively bilingual were separated from those who were native monolingual English-speakers. This was not an independent variable taken into account in the model, so no significant levels are reported.

There was only one native-identified English speaker at UPC, so this participant was not separated within UPC's data. Graphs of the data are present below to aid in the evaluation of hypothesis (4), that learning a second language before the first critical period allows the speaker to fully codeswitch between languages.

Figure (16): Comparing Native Speakers with *Groups*' by *Correlate*

(a) (b)





Surprisingly, the results of the non-native Spanish speakers in W&M (left graph, left grouping) look much more like the data from UPC (right graph, left grouping) than do the results of the native speakers. Because of the high percentage of the W&M speakers that were native Spanish-speakers, the results of the native speakers caused the results of all of the W&M speakers to look more like the native speakers'.

4. Discussion

4.1 Influence of Language

The two *group*s did not differ in their results between their country's dominant *language* and their non-dominant *language*. Therefore, production of native speakers does not align with perception in terms of stress correlates. Speakers, at least those who are native in two languages, have been previously assumed to be able to produce the stress correlates of each language. However, this study shows that perception does not have this same divide; listeners perceive all correlates present no matter the language they hear. Instead, there is likely a correlate perception space that encompasses the correlates of all languages that the speaker knows. Of course, none of the referenced studies have tested L2 speakers' produced stress correlates. There is a small chance, then, that the speakers in the current study would not actually produce stress with different correlates across the two languages. Even if their production space does span both languages, however, the results still point to a space similar to the phonetic one in Flege's Speech Learning Model (2003), which says that bilinguals perceive sounds from all languages in one combined system.

However, the reported discomfort of the UPC participants is not captured by this theory. Hence, the trajectories of improvement as levels increased may also be examined to explain these reactions. The reported unfamiliarity of the UPC participants with regards to the English stimuli was likely due to the lack of schwa in their native phonology, though as bilingual speakers, they were at least slightly accustomed to hearing the vowel. In the English section, the UPC participants had a very steep and linear improvement over the *level*s of the 'combination' *correlate*. However, their results over the *level*s of the other two *correlate*s were much more sporadic. The results for 'combination' therefore reflect higher ability to perceive the *level*s as they increased. Because 'pitch' and 'duration' were similarly fairly flat, the participants must have had more trouble hearing the individual *correlate*s at higher *level*s. Given that this phenomenon only existed in the 'English' section, the factor of *language* may be responsible for the difference in improvements. Therefore, the lack of familiarity with the vowel may have caused the UPC *group* to need both *correlate*s to hear the stress. L2 speakers may therefore use more *correlate*s when perceiving stress in their L2, especially when they are hearing a vowel that does not exist in their L1.

On the other hand, the W&M students were already familiar with the [e] vowel, and therefore may not have had as much need for both *correlate*s to improve. Their results, which showed much more improvement in pitch and the combination than UPC's, may represent more of a RESPONSE to code-switching than to adapting to an unfamiliar vowel; Spanish uses pitch more often when producing stress, so the speakers may have been focusing on pitch to find the location of the stress because the section used Spanish instructions and a vowel extracted from Spanish. English's use of pitch for stress in in-focus words could also play a role in this perception of pitch. Because the vowel does exist in English (the W&M *group*'s country's dominant *language*), the presence of multiple *correlate*s in the Spanish section was not needed as much by W&M as was the presence of multiple *correlate*s in the English section by UPC. When the results of both *language*s are combined, W&M had much stronger improvement in 'duration' than UPC, showing that the lack of improvement in the 'Spanish' 'duration' is a difference from the expected improvement based on the country's dominant *language*.

4.2 Perception Differences in RESPONSE *and Signal Detection Theory*

Relating to *correlate,* the linear model did not find a significant interaction with *group,* though it had an overall significant effect, which will be discussed in section 4.3. Given the use of duration in word-level stress production for 'English' and the use of pitch in word-level stress production for 'Spanish', hypothesis (1) theorized that the W&M speakers would better recognize stress indicated by 'duration' than would UPC speakers, and that UPC speakers would better recognize stress indicated by 'pitch' than would W&M speakers. We do see some support for this difference when looking at slopes and R^2 values.

Based on these results, UPC students were not as able to detect duration-based stress as well as W&M students, and both *group*s were better able to detect the combination *correlate* over the individual ones. The first result supports the hypothesis that production *correlate*s matching the *correlate*s of perception because English (the dominant *language* of W&M) uses duration for word-level stress. The second result affirms the idea that more *correlate*s increase stress perception.

Based on signal detection theory, however, while W&M's sensitivity to 'duration' was much higher than UPC's, it was still not as strong as its sensitivity to 'pitch', whereas based on slopes and R^2 values, it used 'duration' to perceive stress more than it used 'pitch'. Note that because there is no link between the wavelengths of 'duration' and 'pitch' and Hz for each *level*, no absolute comparison can be made. However, the different results from two different tests of the data does prompt a discussion of the two *correlate*s in relation to each other.

If one assumes that there is a connection somewhere between production and perception, the influence of production then may be found in the improvement of perception as *correlate*s' *levels* increase, rather than initial and complete receptiveness to that *correlate* as a whole. This conclusion also applies to the UPC participants: even though signal detection theory points to a very similar perception between 'duration' and 'pitch' for UPC, the *group* had a strong improvement in 'pitch' over the *level*s, more so than 'duration'. The UPC results show a linear increase in 'pitch', rather than the 'duration'. Therefore, once again, the improvement corresponds with the *correlate* of production for the *group*'s dominant *language*[18].

[18] Hypothetically, differing effects of the duration correlate could be linked to Spanish's "syllable-timed" quality. Vogel et al. (2016) stated, "The possibility that syllable-timed languages generally avoid duration as a main stress cue raises an empirical question that needs to

Though signal detection theory typically shows how factors hinder a listener from perceiving differences, the use of this theory in the present study focuses instead of how stress *correlate*s help the listener. Therefore, the ‘combination’ *correlate* helps listeners more than the other *correlate*s individually, whereas in a typical use of signal detection theory, combining factors would cause listeners’ scores to decrease. Instead of participants trying to hear a difference despite factors that hinder perception, the present analysis looks at factors that improve perception.

4.3 Comparing Correlates

When *group* was not broken out, there was a significant difference in RESPONSE between the ‘combination’ and each of the other two. Therefore, listeners have a greater ability to distinguish stress when more *correlate*s are present. This result is not surprising; more assistance in cuing a stressed syllable should cause participants to more easily recognize the location of the stress. The *correlates* ‘duration’ and ‘pitch’ cannot be compared because they are on different scales. The previous pilot studies did establish that for native English speakers, the values for each *correlate* were fairly well-matched, but there is no way to universalize these findings without a real way to establish the starting values and increases among the levels.

4.4 The Role of Position

Both countries had better scores when stress was in ‘antepenultimate’ *position*. This result is surprising given that Spanish has a much higher frequency of penultimate stress than antepenultimate stress, while English does not have such a distinct frequency from that of antepenultimate stress. The participants could theoretically be choosing the less-common option because the study itself does not align with typical *language* environments. The stimuli do not sound completely human, the words are not in the lexicon of either *language*, and the *correlate*s

---

be tested in other languages in future research” (p. 139). However, syllable-timed languages do not cause nearly as rigidly timed syllables as was once thought. Arvaniti (2009) showed that “rhythm” often does not correlate with the actual durations of syllables. Therefore, being syllable-timed or not most likely has no effect on languages’ use of duration as a correlate.

of stress do not all align with those of the participants' *language*s. When the UPC participants hear the stimuli, then, they may simply choose the less natural option to reflect the unnatural environment of the study. Because of the likely influence of the unnatural environment on RESPONSES, the vowel in English may be expected to cause a larger preference for antepenultimate stress for the UPC *group* as they heard a vowel that caused them to find the sentence abnormal. However, the results between the two *language*s of instruction did not differ enough for the vowel to be responsible for the difference. Therefore, the environment, and not the stimuli themselves, likely caused this imbalance.

This theory, however, does not explain the W&M results. Given that four-syllable English words have stress on the penultimate and antepenultimate syllables with equal frequency, there is no reason for the W&M students to choose antepenultimate when the stress was antepenultimate so much more frequently than to choose penultimate when the stress was penultimate. Why were there so many more hits than correct rejections? Here the answer may relate to secondary stress. The stimuli were not given secondary stress in order to focus entirely on the primary stress, but English speakers are accustomed to hearing secondary stress on the first syllable of words with penultimate primary stress. Therefore, the W&M students may have been listening for secondary stress automatically, and, not hearing it, assumed the stress to be antepenultimate. The lack of surface secondary stress for Spanish speakers likely meant that the lack of secondary stress did not affect the UPC speakers.

4.5 REACTIONTIME

Analyzing REACTIONTIME to find the difficulty of certain *correlate*s and *position*s assumes that a longer time to answer is linked with more time required to process the stimuli, which is in turn linked to more effort required to determine the correct answer. Other factors could interrupt this timeline; however limiting the RESPONSE times between .5 and 5s removes most of the answers during which participants got distracted or were not using their time to determine the stress location. Therefore, the results can be assumed to link directly to the difficulty *level*s of the *correlate*s.

The REACTIONTIMES, as shown in section 3.6, were extremely similar to the RESPONSE scores; that is, when the participants more often answered wrong for certain categories of stimuli, they also took longer to answer the same ones correctly. The similarities of these scores show a link, therefore, between incorrectness and processing time required for correct RESPONSES.

4.6 GOODNESS

GOODNESS ratings largely followed RESPONSE scores, so that 'combination' showed higher GOODNESS ratings, just as it had the highest number of correct RESPONSEs. The ratings results did reflect the qualitative comments given by the UPC participants, but not as much as they seemed to think. However, they may have been scaling their RESPONSEs differently for the English and Spanish if they did think that the entire English section was more difficult.

4.5 Native Speakers

The native Spanish speakers from the W&M *group* were not numerous enough to provide much significant information. However, their data is unique from the non-native Spanish speakers and the UPC participants: the RESPONSE scores for the non-native W&M *group* and the UPC *group* followed an increasing line from 'duration' to 'pitch' to 'combination'. Instead, the native speakers at W&M had similar scores for both of the individual *correlate*s (though the score for 'pitch' was slightly lower than 'duration'), with a large increase for 'combination'. There may have been some sort of canceling effect occurring for the stress perception of native speakers. If the results for UPC and W&M are extrapolated to each *group*'s dominant *language*, then Spanish speakers and English speakers must differ their *correlate* preference in stress perception. Therefore, native speakers of Spanish and English speak languages that differ in which *correlate* is most important: the effectiveness of each *correlate* canceled the effectiveness of the other *correlate* out and were not very helpful unless they were used together.

5. Conclusion

The perception of stress position was tested on bilingual Spanish and English speakers at The College of William and Mary in The United States and la Universidad Peruana de Ciencias Aplicadas in Peru. While response scores for the language of instruction (which also correlated with the syllables in each string) did not differ significantly between the two groups, the two did show a difference in their improvement over individual levels: In the 'English' section (the non-dominant language for UPC), UPC did not have as sharp an improvement as W&M did in the 'duration' section, and needed both correlates at once to show the same improvement as W&M. In the 'Spanish' section (the dominant language for UPC), UPC had a steeper slope of improvement for both 'pitch' and the 'combination' correlate. Therefore, participants more easily used correlates in their country's dominant language.

Signal detection theory showed that W&M had a greater sensitivity to 'duration' than 'pitch', but slopes and R^2 values showed that W&M had a greater sensitivity to 'pitch' than 'duration'; however, both methods showed the combination correlate to be the clearest correlate of stress for the W&M group. Signal detection theory did not point to a stronger perception of an individual correlate for UPC, but slopes and R^2 values showed a stronger sensitivity to increases in 'pitch' than 'duration'. Just like W&M, though, UPC most responded to changes in the 'combination' correlate.

Both groups had more correct responses when the stress was in 'antepenultimate' position. For the UPC group, this phenomenon may have occurred because of the perceived strangeness of the task. Spanish has such a preference for 'penultimate' stress that the higher number of 'antepenultimate' responses may reflect a reaction to the words being so unfamiliar; when faced with the task of locating stress, the students chose the more unusual location. For the W&M group, the lack of secondary stress likely caused a preference for 'antepenultimate' primary stress. Words with penultimate stress in English have secondary stress on the first syllable, and stimuli in the task did not have any secondary stress. Therefore, the W&M speakers may have selected 'antepenultimate' responses to reconcile this absence.

REACTIONTIME showed extremely similar scores to the RESPONSE scores, in that the correlates with higher scores also had lower reaction time on the correct responses. GOODNESS also reflected RESPONSE scores in that correlates with higher correct response scores also were given higher goodness ratings, and groups tended to rate their countries' dominant languages with higher goodness ratings.

This study had a relatively small number of participants. A larger scale study along similar lines could prove useful in disambiguating the native speakers of both languages and the L2 learners of one language. Finding a parallel group of bilingual speakers in two separate countries is nearly impossible, and the demographical differences (for example, the high number of native Spanish speakers in the United States when compared with the number of native English speakers in Peru) cause any group to be slightly unbalanced; however, more participants may minimize this unevenness.

In addition, a control study on monolingual English and Spanish speakers could more firmly support the theory of a combined correlate space, assuming that monolingual speakers do not perceive the same correlates as bilinguals. This study worked to establish the differences in stress perception for bilingual speakers in countries with different dominant languages. With further research on teaching the perception and production of specific stress correlates through immersion or explicit instruction, the differences in the trajectory of improvement in the groups of bilingual speakers (as shown in the study) can be minimized in order to better comprehend, and therefore communicate in, the L2.

References

Archibald, J. (1993). The learnability of English metrical parameters by Spanish speakers. *International Review of Applied Linguistics* 31:129-141.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. Phonetica 66:46-63.

Beckman, Mary E. & Edwards, Jan (1994). Articulatory evidence for differentiating stress categories. In Keating, Patricia A. (ed.) *Phonological structure and phonetic form: papers in laboratory phonology III*. Cambridge: Cambridge University Press. 7–33.

Boersma, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.39, retrieved 31 January 2016 from http://www.praat.org/

Bradley, Peter. & Ian Mackenzie. 2004. Spanish: An Essential Grammar. London: Routledge.

Bullock, Barbara E. & Lord, Gillian. 2003. Analogy as a learning tool in the acquisition of L2 phonology. In Ana Teresa P´erez-Leroux & Yves Roberge, *Romance Linguistics Theory and Acquisition*, 281-298. Amsterdam/Philadelphia: John Benjamins.

Clopper, C.G. 2002. Frequency of Stress Patterns in English: A Computational Analysis. Indiana University Linguistics Club Working Papers 2.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech* 45:207–228.

Cutler, A. & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14:113-121. Reprinted in G.T.M. Altmann (Ed.) (2002). *Psycholinguistics: Critical Concepts*. London: Routledge; Vol 1, pp. 157-177.

Cutler, A. (1986). *Forbear* is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, 29:201–220.

Cutler, A., & Pasveer, D. (2006). Explaining cross-linguistic differences in effects of lexical stress on spoken-word recognition. In R. Hoffmann, & H. Mixdorff (Eds.), Speech Prosody 2006. Dresden: TUD press.

Cutler, Anne, and Dennis Norris. "The Role of Strong Syllables in Segmentation for Lexical Access." *Journal of Experimental Psychology: Human Perception and Performance* 14.1 (1988): 113-21. Web. 7 May 2017.

Davis, Colin J., and Manuel Perea. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods* 37.4:665-671.

Delattre, P. (1966). A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics,* 4:183–198.

Dixon, R. M. W. (1988). *A grammar of Boumaa Fijian*. Chicago & London: University of Chicago Press.

Dogil, G., and B. Williams 1999. The phonetic manifestation of word stress. In H. van der Hulst (ed.), *Word prosodic systems in the languages of Europe*, 273-334, Berlin: Mouton de Gruyter.

Dubeda, Tomás & Jan Votrubec. 2005. Acoustic analysis of Czech stress: Intonation, duration, and intensity revisited. In INTERSPEECH. ISCA, 1429 – 1432.

Eriksson, A., Grabe, E., and Traunmueller, H. Perception of syllable prominence by listeners with and without competence in the tested language. *Proceedings Speech Prosody 2002*, Aix-en-Provence, 275-278.

Flege, J. E. 2003. Assessing constraints on second-language segmental production and perception. In Meyer, A. and N. Schiller (eds.), *Phonetics and Phonology in Language Comprehension and Production, Differences and Similarities*. Berlin: Mouton de Gruyter. 319-355.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *JASA* 27:765–768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech* 1:126–152.

Gleitman, L. R., & Newport, E. L. (1995). The invention of language by children: Environmental and biological influences on the acquisition of language. In L. R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science. Language: An invitation to cognitive science* (pp. 1-24). Cambridge, MA, US: The MIT Press.

Goedemans, Rob, and Harry van der Hulst. [2017]. Stresstyp database. http://st2.ullet.net.

Hall, Katherine Lee. 1988. The Morphosyntax of Discourse in De'kwana Carib. Ph.D. thesis, Washington University.

Kim, J.-Y. (2015). Perception and production of Spanish lexical stress by Spanish heritage speakers and English L2 learners of Spanish, In Willis, E., Martín Butragueño, P., &

Herrera Zendejas, E. (eds.), *Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology* (pp. 106-128). Somerville, MA: Cascadilla Proceedings Project.

Lacross, Amy & Liss, Julie & Barragan, Beatriz & Adams, Ashley & Berisha, Visar & Mcauliffe, Megan & Fromont, Robert. (2016). The role of stress and word size in Spanish speech segmentation. *The Journal of the Acoustical Society of America* 140:484-490.

Lehiste, Ilse (1970). *Suprasegmentals*. Cambridge, Mass.: MIT Press.

Lenzo, Kevin. "The CMU Pronouncing Dictionary." *Carnegie Melon University*, http://svn.code.sf.net/p/cmusphinx/code/trunk/cmudict/cmudict-0.7b. Accessed 15 April, 2017.

Llisterri, Joaquim, Machuca, María, de la Mota, Carmen, Riera, Montserrat, & Ríos, Antonio. (2003). The perception of lexical stress in Spanish. In M-J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the XVth International Conference of Phonetic Sciences* (pp. 2023-2026). Barcelona: Causal Productions Pty Ltd.

Lunden, A., Campbell, J., Hutchens, M., & Kalivoda, N. (2017). Vowel-length contrasts and phonetic cues to stress: An investigation of their relation. *Phonology*, 34(3):565-580.

Lunden, Anya & Kalivoda, Nick (2017). Stress correlate database. Available (August 2017) at http://wmpeople.wm.edu/site/page/sllund/stresscorrelatedatabase.

Macmillan, Neil A. (2002). Signal Detection Theory. In Hal Pashler & John Wixted (Eds.), *Stevens' Handbook of Experimental Psychology*: Third Edition, Volume 4: Methodology in Experimental Psychology (43-90). New York: John Wiley & Sons, Inc.

Magen, Harriet. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, 26:381-400.

Martin, Jack B. (2011). *A grammar of Creek (Muskogee)*. Lincoln & London: University of Nebraska Press.

Mattock, Karen, and Denis Burnham. 2006. "Chinese and English Infants' Tone Perception: Evidence for Perceptual Reorganization." *Infancy*, 10: 241–265.

Nadeu, Marianna. 2014. Stress- and speech rate-induced vowel quality variation in Catalan and Spanish. Journal of Phonetics 46:1–22.

Ortega-Llebaria, M., Prieto, P., and Vanrell, M. M. (2008). Perceptual evidence for direct acoustic correlates of stress in Spanish. In J. Trouvain and W. J. Barry

(Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, (pp. 1121–1124). Saarbrücken, Germany.

Ortega-Llebaria, Marta, and Pilar Prieto. 2010. "Acoustic Correlates of Stress in Central Catalan and Castilian Spanish." *Language and Speech*, 54(1): 73–97.

Ortega-Llebaria, Marta, Gu, Hong, & Fan, Jieyu. (2013). English speakers' perception of Spanish lexical stress: Context-driven L2 stress perception. Journal of Phonetics, 41(3-4), 186-197.

Peng, L., and Ann, J. (2001). "Stress and duration in three varieties of English," World Englishes 20:1–27.

Rennison, John R. 1997. Koromfe. Psychology Press.

Saalfeld, Anita. 2012. "Teaching L2 Spanish stress." *Foreign Language Annals*, 45(2): 283–303.

Sluijter, A. M. C., and Heuven, V. J. van (1996). ‘‘Spectral balance as an acoustic correlate of linguistic stress,’’ J. Acoust. Soc. Am. 100:2471– 2485

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language,* 45:412–432.

Taft, L. (1984). *Prosodic Constraints and Lexical Parsing Strategies*, PhD Dissertation, University of Massachusetts.

Tyler, M., Perruchet, P., & Cutler, A. A cross-language comparison of the use of stress in word segmentation. 152nd Meeting, Acoustical Society of America, Honolulu, November 2006. (Abstract: *Journal of the Acoustical Society of America*, 120, 3087).

Vitevitch, Michael S., et al. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and speech* 40.1:47-62.

Vogel, Irene, Athanasopoulou, Angeliki & Pincus, Nadya (2016). Prominence, contrast and the Functional Load Hypothesis: an acoustic investigation.
In Heinz, Jeffrey, Goedemans, Rob & van der Hulst, Harry (eds.) *Dimensions of phonological stress*. Cambridge: Cambridge University Press. 123–167.

Yu, V. Y., and Andruski, J. E. (2010). A cross-language study of perception of lexical stress in English. *J. Psycholinguist. Res.* 39:323–344.

Wang, Q. (2008). *Perception of English Stress by Mandarin Chinese Learners of English: An Acoustic Study.* Doctoral dissertation, University of Victoria, Victoria, BC.

Appendix

A. Generalized Linear Models

		RESPONSE		REACTIONTIME		GOODNESS	
	df	Wald χ^2	*p*	Wald χ^2	*p*	Wald χ^2	*p*
(intercept)	1	159490.232	*p*<.001	24480.155	*p*<.001	42131.602	*p*<.001
group	1	3.255	0.071	41.619	*p*<.001	206.278	*p*<.001
language	1	0.693	0.405	--	--	23.545	*p*<.001
position	1	1352.68	*p*<.001	11.485	0.001	0.358	0.55
correlate	2	11.856	0.003	30.455	*p*<.001	82.209	*p*<.001
level	6	18.13	0.006	34.058	*p*<.001	372.446	*p*<.001
*group*language*	1	0.583	0.445	--	--	2.824	0.093
*group*position*	1	41.649	*p*<.001	0.095	0.758	4.922	0.027
*group*correlate*	2	2.61	0.271	6.616	0.037	3.491	0.175
*group*level*	6	2.34	0.886	7.838	0.025	53.538	*p*<.001
*language*position*	1	5.244	0.022	--	--	3.752	0.053
*language*correlate*	2	1.647	4.39	--	--	2.314	0.314
*language*level*	6	10.7	0.098	--	--	10.646	0.1
*position*correlate*	2	28.903	*p*<.001	3.405	0.182	4.937	0.085
*position*level*	6	135.287	*p*<.001	10.582	0.102	8.989	0.714
*correlate*level*	12	20.167	0.064	18.02	0.115	21.484	0.044
*group*language*position*	1	12.039	0.001	--	--	3.741	0.053
*group*language*correlate*	2	1.888	0.389	--	--	1.526	0.466
*group*position*correlate*	2	1.416	0.493	1.206	0.547	9.884	0.007
*group*language*level*	6	3.193	0.784	--	--	13.701	0.033
*group*position*level*	6	1.498	0.96	11.729	0.068	1.609	0.952
*group*correlate*level*	12	18.087	0.113	8.902	0.711	7.666	0.811
*language*position*correlate*	2	1.069	0.586	--	--	5.697	0.058
*language*position*level*	6	7.767	0.256	--	--	3.205	0.783
*language*correlate*level*	12	11.432	0.492	--	--	16.06	0.189
*position*correlate*level*	12	22.065	0.037	14.565	0.266	10.697	0.555
*group*language*position*correlate*	2	3.043	0.218	--	--	0.542	0.762
*group*language*position*level*	6	5.827	0.443	--	--	1.804	0.937
*group*language*correlate*level*	12	12.944	0.373	--	--	9.422	0.667
*group*position*correlate*level*	12	12.27	0.424	11.771	0.464	6.252	0.903
*language*position*correlate*level*	12	14.639	0.262	--	--	8.734	0.725
*group*language*position*correlate*level*	12	11.249	0.508	--	--	11.05	0.525
participant (group)	26	104.388	*p*<.001	715.132	*p*<.001	914.991	*p*<.001

B. Participant Demographics

						Languages speak:			Native speaker of:		
Participant	Group	Age	Gender	Do you have normal hearing?	Dominant Hand	Spanish	English	Other	Spanish	English	Other
1	W&M	19	F	Yes	L	1	1	French	1	0	
2	W&M	18	F	Yes	R	1	1		1	1	
3	W&M	19	M	Yes	R	1	1	Korean	1	1	Korean
4	W&M	18	M	Yes	L	1	1	French	0	1	
5	W&M	20	F	Yes	R	1	1		0	1	
6	W&M	18	F	Yes	R	1	1	French	1	1	
7	W&M	18	F	Yes	L	1	1		1	1	
8	W&M	18	F	Yes	R	1	1		1	1	
9	W&M	18	F	Yes	R	1	1		1	1	
10	W&M	21	F	Yes	R	1	1	Mandarin Chinese	0	1	
11	W&M	18	F	Yes	R	1	1		1	1	
12	W&M	21	F	Yes	L	1	1	Some Portugese and Italian	0	1	
13	W&M	18	F	Yes	R	1	1		1	0	
14	W&M	20	M	Yes	R	1	1		0	1	
15	UPC	18	F	Yes	R	1	1	French	1	1	
16	UPC	19	F	Yes	R	1	1	Portugese	1	0	
17	UPC	25	F	Yes	R	1	1	French	1	0	
18	UPC	19	M	Yes	R	1	1	French	1	0	
19	UPC	20	F	Yes	R	1	1		1	0	
20	UPC	20	M	Yes	R	1	1		1	0	
21	UPC	18	M	Yes	R	1	1		1	0	
22	UPC	22	M	Yes	R	1	1	Portugese	1	0	
23	UPC	21	F	Yes	R	1	0		1	0	
24	UPC	19	M	Yes	R	1	1	Portugese	1	0	
25	UPC	18	F	Yes	R	1	1		1	0	
26	UPC	19	F	Yes	L	1	1	Portugese	1	0	
27	UPC	20	F	Yes	R	1	0		1	0	
28	UPC	20	F	Yes	R	1	1	French	1	0	

Participant	Group	Competency of country's non-dominant language			Number of years have been learning/speaking country's non-dominant language	
1	W&M	Reading Comprehension	Oral Competency	Listening Comprehension	Age at which learned country's non-dominant language	1 through 3
2	W&M	6	7	6	6 years old or younger	More than 8
3	W&M	6	7	6	6 years old or younger	More than 8
4	W&M	5	6	7	7-12 years old	6 through 8
5	W&M	5	6	5	13-18 years old	3 through 5
6	W&M	6	7	6	7-12 years old	3 through 5
7	W&M	7	7	7	6 years old or younger	More than 8
8	W&M	6	6	6	6 years old or younger	More than 8
9	W&M	5	6	6	6 years old or younger	More than 8
10	W&M	6	6	6	6 years old or younger	More than 8
11	W&M	5	4	4	7-12 years old	More than 8
12	W&M	7	7	7	6 years old or younger	More than 8
13	W&M	7	7	7	7-12 years old	More than 8
14	W&M	5	5	6	6 years old or younger	More than 8
15	UPC	3	5	6	6 years old or younger	More than 8
16	UPC	6	5	6	6 years old or younger	More than 8
17	UPC	6	7	5	7-12 years old	6 through 8
18	UPC	7	7	7	7-12 years old	1 through 3
19	UPC	6	6	6	13-18 years old	6 through 8
20	UPC	6	5	6	13-18 years old	6 through 8
21	UPC	6	5	6	6 years old or younger	More than 8
22	UPC	6	6	6	13-18 years old	6 through 8
23	UPC	5	3	4	7-12 years old	1 through 3
24	UPC	6	6	7	7-12 years old	More than 8
25	UPC	5	5	5	7-12 years old	More than 8

26	UPC	6	4	5	7-12 years old	3 through 5
27	UPC	6	6	6	13-18 years old	6 through 8
28	UPC	6	6	6	6 years old or younger	More than 8

Participant	Group	Did you choose to learn country's non-dominant language?			How did you learn country's non-dominant language?	How often do you converse in country's non-dominant language
1	W&M	Yes	More information	Where did you start learning NON-country language? - Selected Choice	Instruction in country language (non-target language)	
2	W&M	Yes	it was the language spoken at home and the countries i lived in	At home	Complete immersion	Once a day or more
3	W&M	No, I did not learn it on purpose.	my parents raised me speaking Spanish	At home	Complete immersion	Once a day or more
4	W&M	No, I did not learn it on purpose.	I moved to a spanish-speaking country	Elementary school	Complete immersion	At least once a week
5	W&M	Yes		High school	Instruction in non-country language (target language)	At least once a week
6	W&M	Yes		Elementary school	Complete immersion	Once a day or more
7	W&M	Yes		At home	Complete immersion	Once a day or more
8	W&M	No, I did not learn it on purpose.		At home	Complete immersion	Once a day or more
9	W&M	No, I did not learn it on purpose.		At home	Complete immersion	Once a day or more
10	W&M	No, I did not learn it on purpose.	I was taught spanish at an early age but as I grew older I decided to take advance spanish	At home	Complete immersion	Once a day or more

			courses to improve my spanish			
11	W&M	Yes	It was required in middle school to choose to take Spanish or French. I chose Spanish.	Middle school	Instruction in country language (non-target language)	At least once a week
12	W&M	No, I did not learn it on purpose.	It was my first language	At home	Instruction in country language (non-target language)	Once a day or more
13	W&M	Yes	In Kindergarten I was required to take Spanish, all the way through high school, then I chose to study abroad and continue taking Spanish in college	Elementary school	Instruction in country language (non-target language)	Once a day or more
14	W&M	No, I did not learn it on purpose.		At home	Complete immersion	Once a day or more
15	UPC	No, another person decided that I should learn it.	Mi mama me enseno desde que yo era pequena	At home	Instruction in non-country language (target language)	Once a day or more
16	UPC	No, another person decided that I should learn it.		Language school	Instruction in non-country language (target language)	Once a day or more
17	UPC	Yes		At home	Instruction in non-country language (target language)	Once a day or more
18	UPC	No, another person decided that I should learn it.	Mis padres me inscribieron en un instituto	Language school	Instruction in country language (non-target language)	Once a day or more
19	UPC	Yes		Language school	Instruction in country language (non-target language)	Once a day or more
20	UPC	No, another person decided that I should learn it.		At home	Instruction in non-country language	Once a week or more

					(target language)	
21	UPC	Yes		Primary school	Instruction in country language (non-target language)	Once a day or more
22	UPC	Yes		Primary school	Instruction in non-country language (target language)	Once a day or more
23	UPC	Yes		Primary school	Instruction in country language (non-target language)	Once a week or more
24	UPC	Yes		Primary school	Instruction in non-country language (target language)	Once a week or more
25	UPC	No, I did not learn it on purpose.		Language school	Instruction in non-country language (target language)	Once a day or more
26	UPC	Yes		Language school	Instruction in non-country language (target language)	Once a month or more
27	UPC	Yes		Language school	Instruction in non-country language (target language)	Once a week or more
28	UPC		lo aprendi en el nido	Daycare	Instruction in non-country language (target language)	Once a day or more

C. Explanation of Stress/Instructions Script

C1 Spanish Explanation of Stress Script

Antes de la investigación, explicaré el acento prosódico. “El acento prosódico está en la sílaba con la mayor intensidad fonética [(o énfasis)] en la pronunciación”. Las palabras “llame” y “llamé” tienen el acento prosódico en sílabas diferentes. ¿Entiende la diferencia de la intensidad sonora en las sílabas? (PAUSE)

Voy a hacer dos investigaciones hoy día. Le daré las instrucciones para la primera, y luego le ayudaré con la segunda.

Primero, por favor, complete esta actividad. (PAUSE)

C2 English Explanation of Stress Script
I will now repeat the explanation in English. Word stress is on the syllable with the greatest intensity or emphasis in the pronunciation. The words “permit” and “permit” have stress on different syllables. Do you understand the difference in intensity between the syllables?

C3 Spanish Instructions
Escuchará unas secuencias de sílabas repetidas. Estas secuencias no son palabras sino sólo las mismas sílabas repetidas una y otra vez. Cada secuencia tiene cuatro sílabas, y el acento estará en la segunda o la tercera sílaba.

Por ejemplo: ‘aPÓStrofe’ es una palabra real con cuatro sílabas que tiene acento en la segunda sílaba; ‘mariPOsa’ es una palabra real con cuatro sílabas que tiene acento en la tercera sílaba. (PAUSE)

Ud. escuchará algunas muestras de palabras inventadas. Son similares a las secuencias de las investigaciones.

1. Esta secuencia tiene el acento en la segunda sílaba. (PAUSE)
2. Esta secuencia tiene el acento en la tercera sílaba. (PAUSE)
3. Voy a poner dos más: (PAUSE)
4. Como escuchó, el acento estaba en la tercera sílaba para la primera secuencia y en la segunda sílaba para la segunda secuencia. (PAUSE)
5. ¿Quiere escuchar cualquier secuencia una vez más independientemente o al lado de otra? (PAUSE)

Coloque el dedo medio y el índice de la mano izquierda en las teclas ‘2’ y ‘3’ respectivamente.
Si considera que el acento está en la segunda sílaba: Pulse ‘2’
Si considera que el acento está en la tercera sílaba: Pulse ‘3’
Luego, use la mano derecha para hacer clic en un número en la escala al final de la pantalla para indicar la claridad del acento. Haga clic en 1 para indicar poca claridad, y 5 para indicar máxima claridad.

Va a ver una pantalla a la mitad de este proceso que le da un descanso. Cuando quiera continuar, haga clic.

Por favor, llame mi atención cuando termine.

C4 English Instructions
You will hear strings of repeated syllables. Every string has 4 syllables, and either the second syllable or the third syllable will be stressed.

'aSPARagus' is a real 4-syllable word with stress on the second syllable. 'massaCHUsetts' a real 4-syllable word with stress on the third syllable.

You will now hear some samples of four-syllable made-up words that will be similar to what you will hear during the experiment.

1. This word has stress on the second syllable.
2. This word has stress on the third syllable.
3. I'm going to play two more.
4. As you heard, the first string had stress on the third syllable, and the second string had stress on the second syllable.
5. Do you want to hear any of them words again or played next to each other?

Place your left middle and index fingers on keys '2' and '3', respectively.
If the second syllable was stressed, press 2.
If the third syllable was stressed, press 3.
Then, use your right hand to click a number at the bottom of the screen to say how clear the location of stress was. 1 means the least clear and 5 means the most clear.
Click to start.

You'll get a screen halfway through that gives you a break. When you want to continue, click the screen.

Please get my attention when you're done.

D. Reading Comprehension Questions
D1 Spanish
Por favor, lea el párrafo y conteste las preguntas.

"En los días que siguieron, llovió con frecuencia. Para el viernes, cuando el sol finalmente salió, la botellita de aspirinas de Papá estaba vacía y muchas colillas de cigarro cubría el piso del lado de la cama en donde él se acostaba."

Cajas de cartón por Francisco Jiménez, p. 59

1 ¿Cómo se siente Papá?

○ Entusiasmado

○ Alegre

○ Deprimido

○ Contento

2 ¿Cuál palabra describe mejor a Papá?

○ Alcohólico

○ Fumador

○ Estafador

○ Jugador

3 ¿Qué problema probablemente tiene Papá?

○ Dolor de cabeza

○ Fatiga

○ Escalofríos

○ Estornudos

D2 English
Please read the paragraph and answer the questions.

"Estevan left every day around four o'clock to go to work. Often he would come down a little early and we'd chat while he waited for the bus. Everything about him, even his teeth, were so perfect they could have come from a book about the human body. The sleeves of his pressed white shirt were neatly rolled up for a night of dishwashing"

The Bean Trees by Barbara Kingsolver, pg. 123-4

1 How would the narrator describe Estevan?

○ Ugly

○ Stupid

○ Handsome

○ Crazy

2 How does Estevan get to work?

○ Car

○ Public transportation

○ Plane

○ Ferry

3 What is Estevan's job?

- ◯ Washing dishes (1)
- ◯ Teaching classes (2)
- ◯ Fixing cars (3)
- ◯ Directing films (4)