

## Chapter 2

# Segmental cues to IP-initial boundaries: Data from English, Spanish, and Portuguese

© Ricardo Napoleão de Souza

University of Edinburgh

This study investigates the segmental phonetic correlates of IP-initial boundaries in unstressed syllables in three lexical stress languages: English, Spanish, and Portuguese. Using acoustic data gathered under similar experimental conditions, it tests the hypothesis that domain-initial strengthening cues prosodic boundaries in language-specific ways. Moreover, it investigates the role of lexical stress in the phenomenon by focusing on unstressed syllables in post-boundary position, while at the same time testing whether the scope of the strengthening is indeed restricted by measuring segments away from the putative boundary. Results from the analyses of 14 speakers of each of the languages ( $N = 42$ ) strengthen the case for language-specific effects; however, the data suggest that segments farther away from the IP boundary show strengthening as well.

## 1 Introduction

In written language, capitalization signals the beginning of a new sentence, and punctuation marks its end. In speech, various phonetic adjustments play a role similar to capitalization and punctuation, in that they mark the beginnings and endings of prosodic units. In addition to its suprasegmental dimension, a growing body of experimental investigations suggests that prosodic structure influences the phonetic realization of segments in systematic ways (see Fougeron 2001, Cho 2015, for detailed reviews). What's more, the results of these studies suggest that



speakers use language-specific, gradient phonetic detail to group their speech into meaningful units (e.g. Keating et al. 2003, Cho & McQueen 2005).

The investigation of the ways in which prosodic structure modulates phonetic information thus offers important contributions to our understanding of how speech is organized (see also Shattuck-Hufnagel & Turk 1996, Fletcher 2010). Looking at the edges of prosodic units allows speech scientists to determine how the flow of information present in discourse is parsed into cognitively manageable units (Krivokapić 2014). While there is a great deal of research on prosodic domain endings (see Cole 2015 for a comprehensive review), the phonetic encoding of the beginning of prosodic units is much less understood.

There are four main unresolved issues regarding the phonetic manifestation of domain-initial boundaries. First, it is still unclear how boundary-initial prominence relates to other levels of prominence, such as lexical stress (cf. Turk & Shattuck-Hufnagel 2007). Secondly, the evidence on the scope of its temporal effect with regard to the prosodic edge is inconclusive (for a discussion, see Katsika 2016). Thirdly, research has yet to determine the role of language-specific phonology in the phonetic manifestation of domain-initial effects (cf. Cho 2016). Lastly, because most of the literature on domain beginnings has focused on articulation, data on how initial boundaries translate into the *acoustic* signal are still relatively scarce (see Section 1.1 below).

This chapter aims to address these four issues by investigating the acoustic effects of domain-initial IP boundaries on unstressed syllables in prenuclear position in English, Spanish, and Portuguese. The examination of unstressed syllables has the benefit of isolating domain-initial effects from lexical stress, while also allowing for an assessment of their scope of influence. Moreover, a comparison of languages whose unstressed syllables show distinct phonetic properties through similar experimental materials allows for a more straightforward assessment of how language-specific word prosody patterns influence the phenomenon. Since the acoustic results presented here are based on the analysis of 42 speakers of American English, Mexican Spanish and Brazilian Portuguese, results provide more robust quantitative data that may disentangle some of the issues in previous small-scale articulatory studies.

## 1.1 Prosodic boundaries and their markings

One of the central goals of the study of prosody is understanding how speakers group chunks of speech into coherent units according to their meaning and pragmatic functions (cf. Selkirk 1984, Nespor & Vogel 1986, Jun 2006). Although the number of prosodic domains postulated for each language varies, each of

these units is separated from each other through phonetically cued prosodic boundaries (Nespor & Vogel 1986, Shattuck-Hufnagel & Turk 1996, among others). Acoustic correlates of prosodic boundaries include tonal markings, higher intensity and F0 changes, and adjustments in duration (Fletcher 2010). These may differ between initial and final boundaries.

Words that immediately precede a prosodic phrase boundary (i.e. the final edge) have been shown to be consistently longer than those occurring elsewhere in the phrase (e.g. Jun & Beckman 1994, Wightman et al. 1992, Gussenhoven & Rietveld 1992, Berkovits 1994, Byrd et al. 2006, Fougeron 2001, Turk & Shattuck-Hufnagel 2000, Frota & Prieto 2015, among many others). Referred to as *PHRASE-FINAL LENGTHENING*, this type of boundary marking effect has been detected most consistently in the very last syllable of the phrase-final word (cf. Turk & Shattuck-Hufnagel 2007).

On the other hand, boundary marking on the segment near the beginning of prosodic domain has received much less attention in the literature. There is evidence, however, that segments that immediately follow a prosodic boundary (e.g. at the initial edge of an IP) are produced with stronger articulation than when they occur elsewhere in the phrase (cf. Dilley et al. 1996, Fougeron & Keating 1997, Fougeron 1999, Cho & Keating 2001, 2009, Keating et al. 2003, Cho & McQueen 2005, Georgetown & Fougeron 2014, Cho 2011, among others). Because these effects have been noted first in the articulatory realm, boundary-initial marking is most often referred to as *DOMAIN-INITIAL STRENGTHENING* in the literature. Importantly, the articulatory evidence suggests that domain-initial edges are marked solely on the boundary-adjacent segment.

Domain-initial strengthening has been reported in several languages (Cho 2015), although the levels at which it is significant vary depending on the language (e.g. Keating et al. 2003), as well as on individual studies. More specifically, there are reports of segmental correlates of boundary-initial prominence in two of the three languages in the current sample: English (Pierrehumbert & Talkin 1992, Fougeron & Keating 1997, Byrd et al. 2006, *inter alia*) and Spanish (Lavoie 2001, Parrell 2014). However, neither Lavoie (2001) nor Parrell (2014) set out to investigate boundary prominence in Spanish, even though both describe results that could be interpreted as domain-initial strengthening. To the best of my knowledge, no investigation has examined domain-initial strengthening in Portuguese hitherto.

*ACOUSTIC EVIDENCE* for domain-initial strengthening, on the other hand, is much more limited given that it is often reported secondarily in articulatory studies (e.g. Katsika 2016 for Greek; Cho & McQueen 2005 for Dutch; Oller 1973, Tabain 2003 for English; Lavoie 2001, Parrell 2014 for Spanish, Hsu & Jun 1998 for

Taiwanese and Korean). Studies that specifically investigated the acoustic correlates of domain-initial strengthening include Cho et al. (2007), Cole et al. (2007), and White et al. (2020) for English, Italian and Hungarian; Kuzla et al. (2007) for German; and Bodur et al. (2021) for Turkish. Some of the acoustic correlates of domain-initial marking for consonants include voice onset time (VOT), the occurrence of stop burst releases, closure duration, degrees of voicing, duration of nasal murmur, among others. Because individual studies typically investigate one or two variables at a time, it is yet to be determined how different languages use most of those acoustic properties to cue initial edges. Importantly, most of the studies above looked at syllables that were either lexically or phrasally stressed (e.g. under prosodic focus) so that at least some of the effects observed may derive from other types of prosodic elements.

Still, VOT is perhaps the most commonly reported variable connected to domain-initial strengthening. Differences in VOT values correlate with boundary markings on consonants in English (e.g. Pierrehumbert & Talkin 1992, Beckman et al. 1992, Cho & Keating 2001), Dutch (Cho et al. 2007), Korean (Cho & Keating 2001, among others), all of which use VOT to distinguish between /p t k/ and /b d g/ (see Cho et al. 2019 for a discussion). However, VOT is also reported as a cue to initial boundaries in languages with short-lag VOT consonants, such as French (Fougeron 2001) or Japanese (Onaka et al. 2003). The only acoustic correlate of boundary marking on voiceless stops in Spanish is the occurrence of stop burst releases (Lavoie 2001, also for English).

As with the articulatory data, these studies suggest a local effect of domain-initial strengthening, meaning that segmental change due to proximity to the domain-initial boundary occurs on the very first segment following the prosodic edge. On the other hand, studies investigating nasals in nasal consonant-vowel sequences at the IP boundary have found that the vowel in those syllables shows less nasalization (e.g. Cho et al. 2017), which hints at a larger scope of actuation than the initial segment. However, many of the studies reviewed here only investigated individual segments, leaving the question of the scope of the effect somewhat open (see Section 1.3 below).

The body of work presented above suggests that the phonetic variation associated with domain-initial positions correlates systematically with various phrase levels within language-specific prosodic structure. However, a cohesive account of the phenomenon is still lacking despite a growing interest in the so-called prosody-phonetics interface. There are three possible explanations to this observation. First, the number of specificities linked to domain-initial effects reduces comparability between different studies of the phenomenon. Secondly, the interaction of boundary markings with different kinds of prosodic prominence

has often been overlooked in previous research, thereby introducing important confounds. Thirdly, methodological choices in previous studies, including the choice of materials and sample sizes, introduce non-trivial challenges in the interpretation of results. In order to address these issues, the current study uses the same methodology to investigate domain-initial effects in three lexical stress languages: English, Spanish, and Portuguese. The next section addresses key aspects of prosodic structure in each of the three languages.

## 1.2 Prosodic aspects of English, Spanish, and Portuguese

### 1.2.1 Prosodic structure and phrasal prominence

The grouping function of prosody is assumed here to follow a hierarchical structure, meaning that higher levels of structure contain the lower levels, both of which are language-specific (Jun 2014). At the same time, the two highest levels in the prosodic structure, namely the Utterance and the Intonational Phrase (IP), are among the most frequent across languages (cf. Jun 2006, 2014). The IP has been identified specifically as the major prosodic level around which the phonetic correlates of domain-initial boundaries can be measured (cf. Keating et al. 2003).

Indeed, the literature describes the IP as a major domain in English, Spanish, and Portuguese alike (e.g. Beckman & Pierrehumbert 1986, Prieto et al. 1995, Frota 2000). Since there is less agreement for levels below the IP in the three languages, especially regarding mid-level domains (cf. Shattuck-Hufnagel & Turk 1996, Frota & Prieto 2015), this study makes no specific claims regarding levels other than the IP.

English, Spanish, and Portuguese also share other prosodic features. In the three languages, final IP boundaries in declarative utterances are associated with acoustic markings such as final lengthening, pitch declines, and pauses (Beckman & Pierrehumbert 1986, Prieto et al. 1995, Frota 2000). Additionally, the locus of the nuclear accent is similar in the three languages. In neutral declarative sentences, the nuclear accent tends to fall on the rightmost lexical word of the IP and is anchored on the stressed syllable of that word (Pierrehumbert 1980, Hualde & Prieto 2015, Frota & Moraes 2016). Moreover, the nuclear accent can be moved around within the phrase in English, Spanish, and Portuguese, so that any word can potentially receive emphasis (Beckman & Edwards 1994, Ladd 2008, Frota & Prieto 2015, Frota & Moraes 2016; see also Vogel et al. 2018). This prosodic feature common to the three languages was useful in the design of the stimuli for the reading task, as explained in Section 2.2.

On the other hand, the distribution of prenuclear pitch accents differs in English, Spanish, and Portuguese. Spanish and (Brazilian) Portuguese are described as languages with a dense distribution of pitch accents in non-question intonation, so that “essentially every prosodic word (...) receives a pitch accent” (Frota & Prieto 2015: 397).<sup>1</sup> In English, pitch accents other than the nuclear accent are less common than in Romance as a whole, though speakers may accent prenuclear elements (Ladd 2008). Factors influencing the placement of pitch accents in English include semantic-pragmatic factors, structural factors, and rhythmic factors (Shattuck-Hufnagel & Turk 1996).

### 1.2.2 Lexical stress

As hinted above, English, Spanish, and Portuguese all have lexical stress. Lexical stress is obligatory for content words in these languages (Lieberman & Prince 1977, Hualde 2012, Mateus & D’Andrade 2000). Lexical stress is also contrastive in all three (e.g. English [pʰɜːlmɪt] ‘a permit’ vs. [pʰɜːlmɪt] ‘to permit’; Spanish [ˈnumero] ‘number’ vs. [nuˈmero] ‘I number’ vs. [numeˈro] ‘she numbered’; Portuguese [mɛdʒɪkʊ] ‘a doctor’ vs. [mɛˈdʒɪkʊ] ‘I medicate’ vs. [medʒɪˈko] ‘she medicated’).

Stress placement is considered free and difficult to predict in English (e.g. Liberman & Prince 1977). In Spanish and Portuguese, stress follows somewhat more regular patterns. In polysyllabic words, lexical stress generally falls on any one of the three last syllables (Hualde 2012, Mateus & D’Andrade 2000). Although variable, the placement of lexical stress in Spanish and Portuguese can usually be determined based on a series of morphosyntactic and phonological patterns. The main acoustic correlates of lexical stress in the three languages are duration, F0 anchoring, and amplitude (Beckman 1986, Hualde 2012, Mateus & D’Andrade 2000).

Despite the similarities described above, the three languages differ substantially in the degrees to which lexical stress impacts unstressed syllables. Fully unstressed syllables (i.e. not bearing secondary stress) are much shorter than their stressed counterparts in both English and Portuguese (cf. Plag et al. 2011, Cantoni 2013), whereas that difference is less pronounced in Spanish (e.g. Ortega-Llebaria & Prieto 2007). Unstressed consonants show phonetic differences in both English and Spanish, but not in Brazilian Portuguese (Cristófaró-Silva et al. 2019).

---

<sup>1</sup>A prosodic word in Spanish or Portuguese can be defined as a lexical word plus any unstressed clitics (Hualde 2007, Vigário 2003). For instance, the Spanish article *el* in neutral statements such as *el dinosaurio* /eldinoˈsaurjo/ ‘the dinosaur’; or the Portuguese *se* in *feriu-se* /fɨˈɾiʊsɨ/ ‘she hurt herself’. Test words in the current study never contained such clitics and are thus classified as lexical words.

Fully unstressed vowels in English are often centralized to [ə], whereas none of the five Spanish vowels /i e a o u/ shows much qualitative change when unstressed (Nadeu 2014). In Portuguese, fully unstressed vowels show several patterns of reduction, depending on vowel quality, nasality, distance from the stressed syllable and position within the word (Crosswhite 2001, Mateus & D'Andrade 2000, Cristófaró-Silva et al. 2019). For instance, while Câmara JR (1972) described that only five of the seven oral vowels may occur in prestressed position in Brazilian Portuguese (i.e. /i e a o u/ out of /i e ε a ɔ o u/), further variable reduction is now common in those unstressed syllables. Prestressed /i e/ may occur as [i i̯ ɪ̯ ∅], while prestressed /o u/ may appear as [ʊ ʊ̯<sup>w</sup> ∅]. In short, English shows the most consolidated patterns of vowel reduction, whereas Portuguese has been undergoing a series of changes that seem to relate to lexical stress. Spanish unstressed vowels, on the other hand, remain largely unaffected by stress-related reductions. The next section describes how still unresolved issues concerning domain-initial strengthening can be elucidated through an investigation of the languages in the present sample, as well as specific hypotheses and predictions regarding their behavior following IP-initial edges.

### 1.3 Unresolved issues regarding domain-initial strengthening

The three main unresolved issues regarding domain-initial strengthening are their relationship to lexical stress (and other types of prosodic prominence), their scope of influence, and the specific ways in which they interact with language-specific segmental phonology. Another important caveat lies in the fact that most research on domain-initial strengthening has focused on articulatory data from a few speakers. It is possible that inconclusive results regarding those issues could derive in part from experimental design choices and/or from small sample sizes.

While some studies manipulate nuclear accent (e.g. Cho et al. 2017), others may have inadvertently introduced focus accents by having speakers repeat similar or identical carrier sentences (e.g. Lavoie 2001, or Parrell 2014 for Spanish) in which only the test words vary. As a result, these carrier sentences would have likely elicited contrastive focus accents on test words (see Roettger & Gordon 2017 for a discussion). Additionally, a lack of control for other levels of prosodic prominence in test words creates difficulties for the interpretation of experimental findings (cf. Fougeron 2001: 112).

Due to the challenges of collecting articulatory data, many studies have been conducted using small samples of three to five speakers per study. The understandably limited number of speakers in these studies may have nonetheless allowed idiosyncrasies in the speech materials or the behavior of participants to

influence the results. In smaller samples, individual differences in speaker behavior may skew results in ways that make it difficult to distinguish the effects of the variables being tested from those relating to idiosyncrasies in the speech of participants.

More specifically, the current study seeks to provide answers to three research questions (RQ 1–3) stated below.

RQ1: How is domain-initial strengthening expressed acoustically?

While predominantly articulatory in nature, studies of domain-initial strengthening have identified a number of acoustic properties that showed an impact of position within the prosodic domain. Generally, domain-initial strengthening has been found to increase the saliency of segments following the prosodic boundary (Cho 2016), so it is hypothesized that acoustic properties of boundary-adjacent segments will also show an increase in their magnitude. Specifically, the current study is designed to evaluate VOT, burst releases at stop closure, vowel duration, F1 and vowel dispersion.

VOT is the acoustic property that has most often been associated with domain-initial strengthening in several languages (Cho 2016, see also Section 1.1 above). For /p t k/, it is hypothesized that VOT will show greater values following the IP boundary than phrase-medially. Additionally, Lavoie (2001) found differences in the occurrence of stop release bursts in consonants at the beginning versus in the middle of words in both English and Spanish. Consonant bursts are associated with articulatory strengthening (cf. Stevens & Keyser 1989, Torreira & Ernestus 2011). Expanding on the pattern in Lavoie (2001), it is hypothesized that in the current investigation, stops following an IP boundary will show burst releases more often than those occurring IP-medially. Put differently, distance from the prosodic boundary is expected to correlate with more burstless stop releases.

For vowels, Cho & Keating (2009) found partial evidence that domain-initial strengthening increases first formant values in vowels in CV syllables following an IP boundary (see Oh 2021 for similar results for Brazilian Portuguese). Additionally, first formant values serve as an indirect measure of jaw opening, which has been found to correlate with prosodic properties such as prosodic focus (e.g. Erickson 1998). It is expected that vowels showing domain-initial strengthening will thus show higher F1 values than those occurring mid-phrase.

Data on vowel duration is less conclusive, on the other hand. Whereas Oh (2021) found effects of prosodic position within the word (i.e. word-initial vowels in unstressed CV syllables were longer) for Portuguese, target words in her study may have been phrasally accented. Cho & Keating (2009) report no effect



on duration based on proximity to the domain-initial edge, but target syllables in their study received either primary or secondary stress. Although data on vowels in post-boundary CV syllables suggest little to no effect of domain-initial strengthening, potential confounds with lexical and phrasal prominence warrant further testing of its scope (see discussion in Cho 2016).

If the effects of domain-initial strengthening are indeed restricted to the edge-adjacent segment, the vowels in test CV syllables will show no durational differences between prosodic contexts. Alternatively, the hypothesized longer VOT at the IP-initial boundary might increase the overlap between consonant and vowel gestures, leading to shorter and/or devoiced vowels. However, longer VOT at IP-initial CV syllables would still primarily refer to the first segment following the boundary, with any possibly effects on the second segment resulting from assimilation to the former.

RQ2: To what extent is the acoustic manifestation of domain-initial strengthening manifested language-specifically?

Cho & McQueen (2005) and Cho (2016) discuss how different languages show specific combinations of acoustic cues to boundary marking. According to Cho, segments subject to domain-initial strengthening are “fine-tuned (...) making reference to the phonetic content provided by the language-specific phonetic feature system” (Cho 2016: 136). In other words, the phonetic expression of initial boundary marking depends on which features the language already uses to convey phonological distinctions such as /k/ vs. /g/, or /ʊ/ vs. /u/. The current study uses VOT and differences in unstressed vowel qualities to test Cho’s hypothesis.

For instance, VOT serves different roles in English vs. Spanish or Portuguese. In English, /p t k/ are distinguished from /b d g/ mostly through VOT, whereas Spanish and Portuguese primarily use voicing to signal the same distinction. Recent data on Brazilian Portuguese, however, suggest that longer VOT (“aspiration”) may be emerging as an acoustic cue to /p t k/, albeit with smaller values than English (cf. Alves et al. 2008, Ahn 2018). Based on these patterns, it is expected that domain-initial strengthening would be observed in the current sample through VOT by showing the largest differences between prosodic contexts in English, followed by Portuguese, and the smallest differences in Spanish (see Figure 1).

RQ3: How does domain-initial strengthening affect unstressed syllables in pre-nuclear position within the Intonational Phrase?

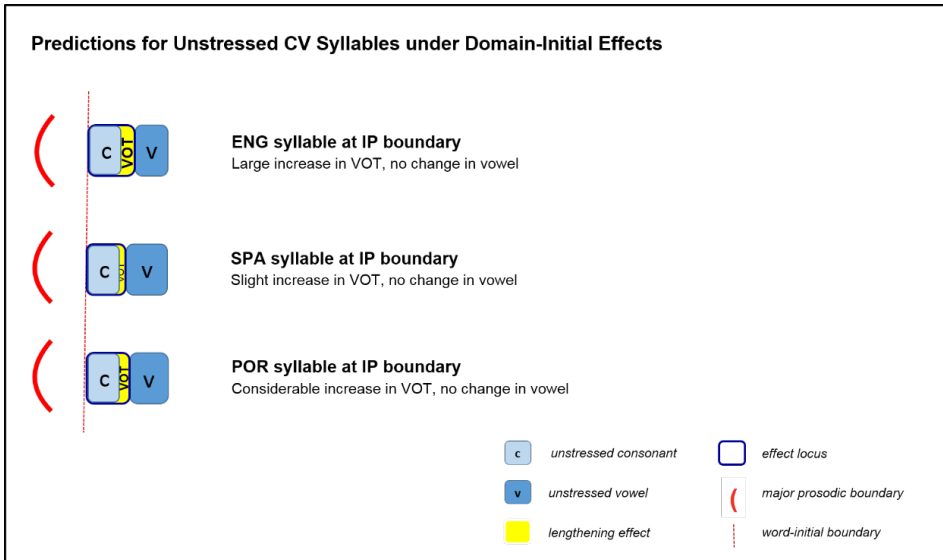


Figure 1: Schematic representation of the expected effects of domain-initial strengthening on unstressed syllables at an IP boundary in the languages in the sample, based on previous findings.

As reviewed above, a great number of studies investigating domain-initial strengthening failed to control for lexical or phrasal prominence (but see Kim et al. 2018). As such, it remains to be determined how the marking of initial prosodic boundaries influences the segmental makeup of unstressed syllables in prenuclear position. One general hypothesis guiding the current investigation is that any effects of boundary marking would be more apparent in those syllables.

If the findings in previous research hold, one would observe boundary-induced changes only to the segment following the IP edge, for instance the consonant in a CV syllable (see Figure 1). The vowel in such syllables would thus manifest the same language-specific characteristics of unstressed vowels near the IP edge as in the middle of it, for instance shorter duration and/or centralization. As mentioned above, the interplay of stress-related reduction on the vowel and boundary-induced strengthening on the consonant in languages like English or Portuguese could potentially lead to changes in the syllable itself, for instance through emerging devoiced vowels due to increased gestural overlap (cf. Jun & Beckman 1994, Mo 2007 for Korean, Davidson 2006 for English, Delforge 2008 for Spanish).

In sum, the current study expands on previous findings by addressing the following key points:

1. Separating domain-initial effects from lexical prominence by investigating word-initial unstressed syllables.
2. Controlling the placement of nuclear accent in the target sentences through elicitation of narrow focus away from the initial boundary.
3. Avoiding an excessive influence of individual idiosyncrasies on the overall results by recording a larger group of speakers per language.
4. Evaluating the role of language-specific phonology by performing a direct comparison of three languages using similar materials.

## 2 Experimental materials and methods

This study investigated domain-initial effects on fully unstressed syllables in pre-nuclear words in three languages with contrastive lexical stress. The method presented here focuses on speech data sampled under experimental control. Separate reading tasks were conducted with native speakers of American English, Mexican Spanish, and Brazilian Portuguese using the same experimental procedure and analysis methods.

### 2.1 Participants and experiment procedure

In total, 42 participants, aged between 18 and 31, took part in the reading experiment. Fourteen speakers of each of the three languages make up the sample (English: eight female; Spanish: twelve female; Portuguese: eight female). Participants were university students, mostly undergraduate, and were naïve to the purposes of the study. All participants were native L1 speakers of their respective sample language, with no known vision, hearing, or speech impediments.

All English speakers reported being monolingual; Portuguese speakers were bilingual in English to different degrees. All Spanish speakers were fully bilingual in English, having nonetheless first acquired Spanish in the home from both of their (Mexico-born) parents. Most speakers of each language came from the same state within their respective countries (American English: New Mexico; Mexican Spanish: Chihuahua; Brazilian Portuguese: Minas Gerais). However, no specific efforts were made to control for dialect representation within country varieties, so that any dialectal differences that may have arisen are not accounted for here.

Experiments took place in soundproof or quiet rooms at the University of New Mexico in the United States, and at the Universidade Federal de Minas Gerais

in Brazil. The full reading session lasted between 25 and 50 minutes. Acoustic data were acquired at a sampling rate of 16 kHz through an Audio-Technica USB microphone plugged directly into a laptop. The sound editing software Audacity (Audacity Team 2014) was used to record all participants.

Stimuli were presented on sheets of paper with six to eight sentences per page in large fonts, one sentence per line. Participants read each carrier sentence three times in pseudo-randomized order, which were interspersed among filler items eliciting various other types of prosody (e.g. lists, questions, etc.). Additionally, different presentation orders were devised for each language, so that the items were read in the same order by only three or four speakers per language. Reading sessions were divided into two (for English) or three blocks (for Spanish and Portuguese) to avoid experimental fatigue. Speakers were encouraged to read the sentences at their own pace.

English speakers read 120 sentences in total whereas Spanish and Portuguese speakers read 180 each. These included sentences with test words beginning with /m n/ which were intended for future analysis. A short practice run was conducted before the recording started. Unless prompted by participants themselves, the experimenter provided no feedback/corrections during recording sessions. However, whenever there was an error or disfluency, speakers could repeat a given sentence if they wished. The experimenter interacted with participants in their own respective languages.

## **2.2 Experiment materials**

### **2.2.1 Target syllables and test words**

Target syllables consisted of a CV sequence of /p t k/ plus a monophthong in unstressed, word-initial position. For English, the vowel in the nucleus of the target syllable was always /ə/, selected to avoid possible confounds of secondary stress (cf. Crystal & House 1988, de Jong 2004). In both Spanish and Portuguese, target syllables had high and low vowels (i.e. /i a u/). The use of two different high vowels in the Romance languages was due to language-specific constraints in the distribution of stop plus high vowel in those languages. Vowel height was then included in models as dependent variables due to differences in duration in high and low vowels in both Spanish and Portuguese (cf. Hualde & Nadeu 2014, Cristófar-Silva et al. 2019). Test items were trisyllabic words with penultimate stress in all three languages. Using three-syllable words reduces the possibility of secondary stress due to rhythmic constraints. Most test words had the overall /CV'CV.CV/ shape. Immediately following target syllables in all test words was

the stressed syllable, which always started with a voiceless stop or a voiceless affricate. A total of 6 English words, and 12 Spanish and 12 Portuguese words were tested in this investigation. Table 1 illustrates the test words used in the experiment by the consonant in the word-initial syllable (see Appendices A–C for the full list).

Table 1: Examples of target words used in the English, Spanish, and Portuguese stimuli. Note that the sequence [ti] is absent in most dialects of Brazilian Portuguese

Language	Segment		
	/p/	/t/	/k/
English	p <sup>h</sup> ə'tɪʃənz 'petitions'	t <sup>h</sup> ə'kɪlə 'tequila'	k <sup>h</sup> ə'pɛɪfəz 'capacious'
Spanish	pi'kete 'injection prick'	ti'pexo 'an idiot'	ki'tepo 'person from Quito'
Portuguese	pɾɨ'tadə 'a pinch of'	tɔ'telə 'guardianship'	kɾɨ'tadə 'paid-off'

### 2.2.2 Experimental conditions and carrier sentences

This study uses the IP as test ground for its hypotheses. Test words were tested in two different conditions in carrier sentences: either at the beginning of an IP (IP-initial condition) or in the middle of the IP (IP-medial condition).<sup>2</sup> Thus, word-initial target syllables were themselves either IP-initial or IP-medial. Test words always occurred in prenuclear position. Overall carrier sentence size was kept at 25 canonical syllables in each of the three languages, for both conditions. This precaution was taken as a control for utterance length effects on articulation rate, and thus segment duration (cf. Fónagy & Magdics 1960, among others).

Following Cho & McQueen (2005), the IP containing the target syllable was always preceded by a precursor IP to ensure that the target syllable necessarily occurred at the beginning of the IP domain rather than at the Utterance. In turn, the position of the target syllable within the test IP was kept constant for each

<sup>2</sup>One could alternatively call the two conditions “IP-initial” and “Wd-initial”. However, given that target syllables were always word-initial, the “Wd-initial” condition is referred to as “IP-medial” throughout the chapter to avoid confusion.

condition across languages. In the Spanish and Portuguese stimuli, the target syllable always occupied the seventh slot in the IP-initial condition, whereas in the English sentences it was always the sixteenth in the carrier utterance. Punctuation marks (i.e. colons, semicolons, or commas) elicited a separation between the precursor and test IPs (cf. Turk & Sawusch 1997, Keating 2004).

A further series of measures were undertaken to guarantee that test words did not receive nuclear pitch accent. This is because in all three languages, nuclear pitch accent has been found to increase duration in words/syllables that bear the main IP accent (e.g. Cantoni 2013 for Brazilian Portuguese; Hualde & Prieto 2015 for Spanish). Contrastive or corrective narrow focus was thus elicited elsewhere in the IP in the IP-initial condition. In the three languages, focused words receive the nuclear pitch accent.

In each carrier sentence in the IP-initial condition, the precursor IP(s) established the context for narrow focus in the IP that contained the test words. The exact number of precursor IPs or words in them varied due to differences in word size and syntactic factors between the languages. In addition to contextual information, focused words were capitalized and participants were instructed to place emphasis on them (see Turk et al. 2006 for a discussion of this strategy). Table 2 illustrates the carrier sentences used in the study (see Appendices A–C for the full set of stimuli, with translations).

Additionally, carrier sentences consisted of a variety of meaningful passages, so as to avoid inadvertently eliciting contrastive focus on test items which occurs in sentences of the “Say x again” type. Finally, in order to avoid monotonous intonation due to experimental fatigue (Xu 2010), filler items consisted of a variety of sentence types of different sizes, including yes-no and wh-questions, lists, exclamations, and contrastive focus sentences (see Appendices A–C). All filler sentences were obtained from the Brigham Young University (BYU) corpora of the specific varieties of the languages investigated: *Corpus do português* (Davies & Ferreira 2006), *Corpus del español* (Davies 2016), and the *Corpus of Contemporary American English* (Davies 2008).

### **2.3 Variables**

The acoustic marking of domain-initial position was evaluated separately for each segment, so that there are two broad types of dependent variables: consonant measures, and vowel measures. Independent control variables include phonological as well as sociolinguistic variables (i.e. age, gender). These are described in turn below.

Table 2: Sample carrier sentences for test words ‘capacious’, *pitada* ‘whistle blow’/‘pinch (e.g. of pepper)’ in Spanish and Portuguese, respectively. The sentences below show how unstressed CV syllables (e.g. *kə* in ‘capacious’ [k<sup>h</sup>əp<sup>h</sup>ɛɪfəz]) were tested under the two experimental conditions in this study. The IP-initial boundary is represented by <||>; the IP-medial boundary is represented by <#>

Cond.	Lang.	Carrier sentence
IP-initial		
	ENG	It doesn’t refer to ability! You can check for yourself    capacious means ROOMY or full of space
	SPA	‘Estás confundido    pitada quiere decir SOPLADO más que sonido o pitido
	POR	Tem muito sal aqui,    pitada quer dizer só um POUQUINHO do ingrediente na receita
IP-medial		
	ENG	It is very sad there’s not too much they can do at this point: the city’s # capacious museum closed
	SPA	A causa de la lluvia, el árbitro Federico dió la # pitada a las tres horas
	POR	Pimenta caiena é mais forte do que do reino, só uma # pitada tá mais que bom

### 2.3.1 Variables pertaining to /p t k/ in target syllables

#### 2.3.1.1 Voice onset time (VOT)

Defined here as the interval immediately after the release of the stop up to the onset of voicing (Lisker & Abramson 1964). VOT was measured manually from first peak of the stop burst release up to the zero crossing nearest the onset of the second formant in the following vowel, as shown in the spectrogram. In case of multiple release bursts, the first burst was used. In the absence of clear burst releases, periods of visible aspiration in the spectrogram were also measured as VOT (cf. Abramson & Whalen 2017), in which case the beginning of the aspiration noise was taken as the acoustic delimiter for segmentation.

#### 2.3.1.2 Occurrence of stop release burst(s)

Defined here as a transient noise pulse at the release of the built-up air pressure during the voiceless stop closure. The occurrence of a stop release burst was measured by visual inspection of both waveform and spectrogram.

### **2.3.2 Variables pertaining to /ə/ or /i a u/ in target syllables**

#### **2.3.2.1 Vowel duration**

Vowel duration was measured from the onset of the vowel's F2 to its offset as seen in the spectrogram and the waveforms.

#### **2.3.2.2 First formant (F1)**

F1 was extracted from the midpoint of the target vowel as labeled in the TextGrid. F1 data were subsequently normalized for each token based on the means and standard deviations calculated over all productions by the same speaker.

#### **2.3.2.3 Fundamental frequency (F0)**

Fundamental frequency may serve as a correlate of lexical prominence. For all tokens produced by female speakers, the range of analysis for F0 was specified as 100–400 Hz. Tokens from male speakers were analyzed using a range of 50–250 Hz. If domain-initial strengthening shows no influence on F0, values would not differ significantly between conditions.

#### **2.3.2.4 Vowel dispersion**

Vowel dispersion was measured only in the Spanish and Portuguese data. It is defined as the location of vowels along the back-front and high-low dimensions measured as a function of F1 and F2. As with F1, F2 was extracted from the midpoint of the vowel as labeled in the TextGrid.

### **2.3.3 Independent control variables**

#### **2.3.3.1 Silent interval**

Silent interval is defined here as the interval without vocal fold vibration in the waveform, in milliseconds. For sentences in the IP-initial condition, silent interval duration was measured from the end of the last segment of the last word in the precursor IP up to the beginning of the first segment of the test word (as generated by the automatic forced aligner). Silent interval duration was measured in the IP-medial condition using the same criteria as in the IP-initial condition. It should be noted that the automatic forced aligners might include part of the closure of the voiceless stop in its measure of the putative pause that occurs before the test word.



### 2.3.3.2 Duration of the stressed syllable in the test word

The stressed syllable is defined as the most prominent syllable in a word, as specified in the lexicon of each of the three languages. In the Spanish and Portuguese data, stressed syllables were measured based on the output of the purpose-designed Praat script. In the English data, the stressed syllable was measured manually from the F2 offset of the unstressed vowel in the target syllable up to the offset of F2 in the vowel in stressed syllable for all but one test word. In the case of the word ‘patrolmen’, the only test word whose stressed syllable ended in a coda consonant, the syllable was measured from the F2 offset of the unstressed vowel /ə/ up to the beginning of the nasal murmur of the /m/ in the final syllable /mn/. Stressed syllable duration values were log-transformed for analysis given differences in the number and types of segments in the stressed syllable across the different words used as stimuli (see Section 2.2).

### 2.3.3.3 Articulation rate

Defined as the number of syllables divided by phonation time, excluding silent intervals over 200 milliseconds. Articulation rate was measured automatically through a Praat script (de Jong & Wempe 2009), which estimated the number of syllables based on acoustic information in the audio files containing individual carrier sentences. The script identifies syllable nuclei by detecting peaks in intensity (dB) that occur between two dips in the audio file, thus avoiding measuring segments other than vowels. It is assumed that the faster the articulation rate, the shorter the acoustic durations (cf. Fónagy & Magdics 1960, Crystal & House 1988, see also Kessinger & Blumstein 1998 specifically for VOT).

### 2.3.3.4 Place of articulation

Place of articulation has been shown to influence VOT values (Cho & Ladefoged 1999, for crosslinguistic data; Avelino 2018, for Mexican Spanish; and Ahn 2018, for Brazilian Portuguese), as well as stop burst release (Winitz et al. 1972).

### 2.3.3.5 Vowel height

Vowel height constituted a variable only in the analyses of Spanish and Portuguese. Vowel height was coded as high, or low, based on citation forms of the vowels in target syllables. High vowels were /i u/, and the low vowel consisted of /a/ alone.

#### 2.3.3.6 Repetition

Each speaker read the stimuli three times. This variable identifies the order of the individual productions of each target word in the reading task: first, second, or third, for each speaker. Previous results (e.g. Fowler & Housum 1987) suggest that segment duration in the second or third productions will be shorter than the first occurrence in the stimuli.

#### 2.3.3.7 Word frequency

Word frequency was operationalized as the number of occurrences of target words per million in the Corpus of Contemporary American English (Davies 2008), the Corpus del español (Davies 2016) for Spanish, and the Corpus do português (Davies & Ferreira 2006), for Portuguese. These values were log-transformed for the statistical analysis. Frequency of occurrence of a lexical item correlates with its overall duration (Bell et al. 2009), so that the higher the frequency, the shorter the duration.

### 2.4 Specific criteria for confirming IP-initial boundaries

This study compares words produced at IP-initial boundary against those produced phrase-medially. Hence, it was crucial that the production of carrier sentences matched the prosodic context they were designed to elicit. The presence of a long silent interval (i.e. 200 ms or more) between the precursor and the test IP was used as the primary criterion for determining the occurrence of a prosodic boundary. Silent intervals are particularly relevant in the current study since they can serve as indications of prosodic boundary strength (see Krivokapić 2014 for a review). Specifically, there is evidence to suggest that the duration of a pre-boundary pause may correlate with gestural magnitude of the first segment following the pause (e.g. Beňuš & Šimko 2014, see also Ramanarayanan et al. 2009).

Whenever the silent interval between precursor and test IP was shorter than 200 milliseconds, two additional acoustic parameters were used as a confirmation that test words in the IP-initial condition were in fact produced at the left edge of the phrase: pitch declination and reset, and/or the presence of creaky voice (i.e. “phrase-final creak”, see Garellek 2015) in the last word in the precursor IP. Both parameters were assessed by visual inspection of the waveforms and spectrograms.

Pitch declination and reset is defined as a lowering of the pitch range between the early part of the precursor IP and the end of that IP, without regard to the

tonal description (e.g. HL% or L%). Such lowerings were always followed by a reset of the pitch excursion, meaning that the speaker reset their pitch at the start of the test IP at a higher level than that of the precursor IP. Pitch declination and reset have been described as a cue to IP boundaries in broad statements in all three languages in the sample (Pierrehumbert & Hirschberg 1990, for English; Frota et al. 2007, for Spanish and Portuguese).

In turn, phrase-final creak is defined as a stretch of the speech signal characterized by irregular (e.g. less periodic) F0 and amplitude changes (Redi & Shattuck-Hufnagel 2001, Garellek 2015) that occurs at or near the end of a prosodic phrase. Phrase-final creak has been found to correlate with the end of larger prosodic domains in English (Redi & Shattuck-Hufnagel 2001), Spanish (de la Mota et al. 2010), and Portuguese (Frota & Moraes 2016) alike. Creaky voice was noted as phrase-final creak when it affected all voiced segments in the last word of the background sentence (Garellek 2015) for English and at least the last syllable for Spanish and Portuguese.

## 2.5 Data extraction, data exclusion

Acoustic measurements were done in Praat (Boersma & Weenink 2019). The FAVE-align automatic forced aligner (Rosenfelder et al. 2014) was used to generate segment intervals in TextGrids for the analysis of the English stimuli, which were then hand-corrected as needed. Syllable, word, and phrase intervals were created manually based on the output for segments generated by FAVE. For the Spanish and Portuguese datasets, the automatic forced aligner EasyAlign (Goldman 2011) generated syllable, word, and phone intervals. Hand-corrections were made where needed. Subphonemic segmentation was done manually based on the acoustic information available in the spectrogram and waveform. Prosodic annotations were also done manually. All annotated data were extracted automatically from the TextGrids using purpose-designed Praat scripts.

Each speaker produced 36 tokens of the test words (6 words  $\times$  2 conditions  $\times$  3 repetitions), totaling 504 tokens per language prior to inspection. Tokens produced with unexpected prosodic or intonational patterns, such as laughter, hesitancy, or misplaced nuclear accent on test words, were excluded from analysis. Errors affecting the precursor IP also led to exclusion, although sentences containing errors affecting words that followed the test word were kept. Finally, test words in English or Portuguese showing vowel deletion in the target syllable were also excluded from the acoustic analyses (50 English tokens, and 31 Portuguese tokens). In total, 423 tokens of English data were analyzed. For Spanish, the data correspond to 413 tokens, whereas Portuguese results derive from 400 tokens.

## 2.6 Statistical analyses

All statistical analyses were conducted in R (R Core Team 2018). Visual inspection of the data involved generating basic graphs that displayed broad distribution patterns of the dependent variables across experimental conditions. The Shapiro-Wilk normality test was then applied to variables using the generic function built in R as a way to assess whether values in numeric variables followed a normal distribution. For instance, the Shapiro-Wilk test confirmed that VOT, unstressed vowel duration, and F1 data failed to reach a normal distribution in all three languages. The two-tailed Wilcoxon non-parametric statistical hypothesis test was then applied separately to the three continuous dependent variables in each language to determine whether prosodic context yielded statistically significant differences in the data (see Wilcoxon tests results). Variables that showed no statistical difference between conditions were excluded from further analyses of the individual languages (see Results). Mixed effects models were only fit for variables that showed statistically significant differences between experimental conditions.

Numeric variables such as duration and VOT were log-transformed. Additionally, numeric predictors were z-scaled using the generic function `scale()` in R (cf. Gries 2013, Bell et al. 2009). At this point, any remaining outliers (i.e. figures that were three median absolute deviations away from the median) were further excluded from the subset. Linear mixed-effects models were then fit to each subset of the data using the `mixed()` function in the Afex package (Singmann et al. 2015) with all appropriate independent variables as predictors. The `mixed()` function also produces *p*-values for the likelihood ratio test.

Variables in each model were introduced through a backward selection procedure to help guard against model overfitting. Following this procedure, the first model was fit with all individual predictors and theoretically relevant interactions. Interactions tested included: articulation rate vs. repetition, duration of the stressed syllable vs. condition, place of articulation vs. VOT, vowel height vs. vowel duration (for Spanish and Portuguese only), and vowel duration vs. repetition. The interaction of duration of the stressed syllable and condition specifically tested whether proximity to the IP-initial boundary influenced stressed syllable duration. After each model was fit, it was compared to a set of models with one fewer predictor via the generic function `anova()` in R. The Akaike Information Criterion (AIC) was used as a goodness-of-fit measure for model comparison. The predictor that contributed the least to model fit was then removed from the full model. The process was repeated until the final model was significantly better than all possible alternatives with one fewer predictor.

### 3 Results

This study investigated how the marking of domain-initial boundaries affects the acoustic properties of segments in fully unstressed syllables in three lexical stress languages. The following results stem from comparing test words measured under two experimental conditions: IP-initial, and IP-medial. Results are presented first as an overview, followed by detailed descriptions of the individual languages (see also Statistical summaries and Wilcoxon tests results for statistical data).

#### 3.1 Crosslinguistic summary

Table 3 summarizes the results of the comparison between prosodic contexts in the three languages in the sample. All languages showed differences between IP-initial and IP-medial contexts, although the specific acoustic correlates varied somewhat between the languages.

Table 3: Comparison of results by prosodic condition in the three languages (significance levels are \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ ; n.s.: not significant; IP: IP-initial; Wd: Word or IP-medial).

Measure/Language	English	Spanish	Portuguese
VOT /p t k/	IP > Wd *	n.s.	n.s.
Vowel duration	n.s.	IP > Wd ***	IP > Wd ***
F1	n.s.	n.s.	n.s.
F0	n.s.	IP > Wd ***	n.s.
Vowel dispersion	NA	n.s.	n.s.
Burst release	IP > Wd **	IP > Wd ***	IP > Wd **
Stressed syllable duration	IP < Wd ***	IP < Wd ***	IP > Wd ***

As shown in Table 3, the languages in the present study show some similarities and quite a few differences in how boundary-initial marking affects the acoustic properties of trisyllabic words with penultimate stress. All languages showed more burst releases at stop closure for /p t k/ that followed the IP-initial boundary than IP-medially. On the other hand, VOT is used to mark stops in unstressed syllables only in English, with Spanish and Portuguese /p t k/ failing to show significance in how VOT lag differs between prosodic contexts (see Figure 2). Additionally, whereas tokens of English /ə/ in the target CV syllable were unaffected by boundary marking, the duration of Spanish and Portuguese /i a u/ under similar conditions shows significant effects of prosodic context (Figure 3). Spanish words, in particular, also showed higher F0 values near the IP boundary.

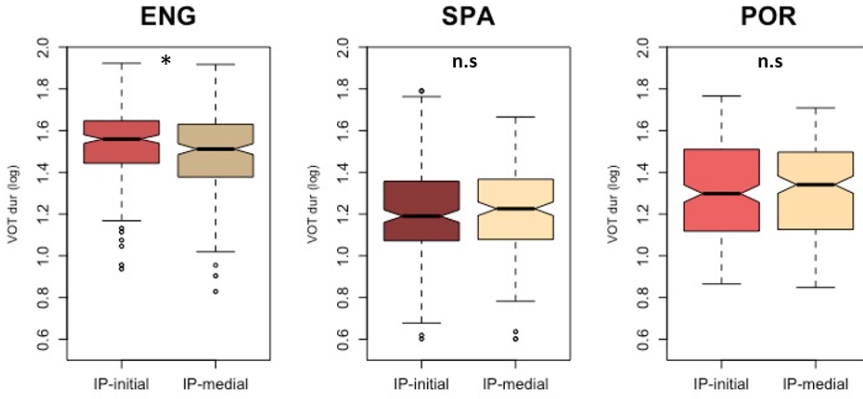


Figure 2: Duration of VOT lags in target syllables (log values). Box plots show VOT duration results measured in target syllables across the stimuli (e.g. English “capacious”, Spanish/Portuguese *pitada*), as a function of prosodic context. IP-initial values in left box, IP-medial data in right box; y-axis shows the same scale for the three languages; \* $p < .05$ ; n.s.: not significant.

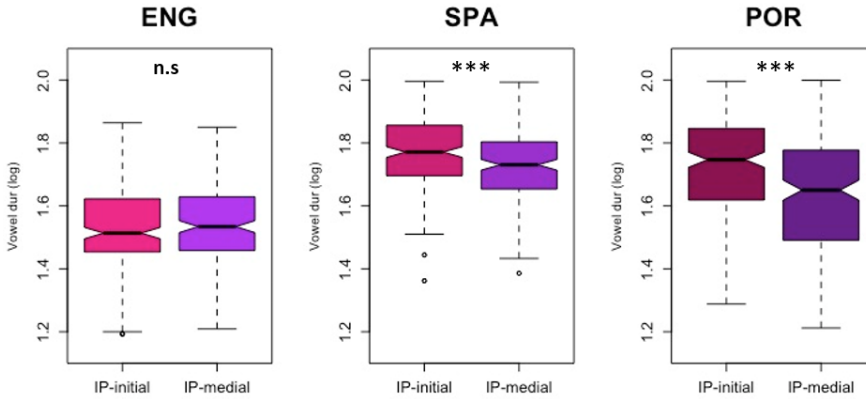


Figure 3: Duration of vowels in target syllables (log values). Box plots show vowel duration results measured in target syllables across the stimuli (e.g. English “capacious”, Spanish/Portuguese *pitada*), as a function of prosodic context. IP-initial values in left box, IP-medial data in right box; y-axis shows the same scale for the three languages; \*\*\* $p < .001$ ; n.s.: not significant.

What's more, while the non-boundary adjacent stressed syllable shows systematic variation in all three languages, such an effect is uneven; English and Spanish words pattern in one way, and the Portuguese data in another. In both English and Spanish, the stressed syllable was shorter near the IP boundary, whereas in Portuguese it was longer. It should be noted, however, that stressed syllables were control variables, so that their phonetic content in the current data was not controlled for. The following sections detail these findings in light of the statistical techniques used to evaluate them.

### 3.2 Consonant results: unstressed /p t k/ in target syllables

#### 3.2.1 Burst release at stop closure

The occurrence of a burst release at the stop closure was evaluated for each language via Fisher's exact tests. Based on Lavoie (2001), it was expected that consonants at an IP-initial boundary would show more burst releases than those occurring IP-medially. This prediction was borne out in the data, although it is noteworthy that there was an overall high rate of burst releases for all consonants in the sample. Table 4 shows the burst release results for the three languages.

Table 4: Proportions of occurrence of burst at /p t k/ release in the sample by prosodic context;  $N$  = number of tokens. Significance levels are \*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$

	English**		Spanish***		Portuguese**	
	Yes (N)	No (N)	Yes (N)	No (N)	Yes (N)	No (N)
IP-initial	1.0 (222)	0	1.0 (221)	0	1.0	0
IP-medial	0.95 (200)	0.05 (10)	0.94 (11)	0.03 (11)	0.97 (176)	0.03 (6)

#### 3.2.2 VOT in English

VOT values were evaluated through mixed effects models with word, and speaker as random effects. As mentioned above, English was the only language in the current study that showed statistically significant differences for VOT values in target syllables between prosodic conditions. A total of 432 observations related to VOT in voiceless stops were entered in the English mixed-effects model. The model was fit through backward selection of variables, meaning that predictors that failed to improve model fit (i.e.  $p \geq 0.05$ ) were excluded from the analysis.

The best model for VOT (log-transformed) as the response variable included speaker and word as random effects, and the following variables as fixed effects: prosodic context (two levels: IP-initial or IP-medial), consonant type (three levels: /p/, /t/, or /k/), and the interaction of duration of the stressed syllable and prosodic context. The R function `mixed()` automatically calculated *p*-values. The final fixed and random effects estimates appear in Tables 5 and 6, respectively.

Table 5: Main effects for the English model with VOT in unstressed word-initial /p t k/ as the response variable (log-transformed). The reference levels for categorical predictors are IP-initial for prosodic context and ‘/k/’ for consonant type. Significance levels are \*\*\**p* < 0.001; \*\**p* < 0.01; \**p* > 0.01; n.s.: not significant.

Estimate	$\beta$	SE	<i>t</i>	<i>p</i> ( <i>t</i> )
Intercept	1.900	0.397	4.78	0.004**
Prosodic context = IP-medial	−1.831	0.588	−3.11	0.002**
Stressed syllable duration x IP-medial	0.642	0.214	3.00	0.009**
Consonant = /p/	−0.247	0.028	−8.80	0.166 n.s.
Consonant = /t/	−0.059	0.028	−2.11	0.389 n.s.
Stressed syllable duration x IP-initial	−0.127	0.172	−0.74	0.492 n.s.

Table 6: Random effects intercept – English VOT in target syllables (SD: standard deviation)

Variable	Variance	SD
Speaker	0.0088	0.0941
Word	0.0002	0.0126
Residual	0.0445	0.2110

As summarized in Table 5, the results from the generalized linear model revealed significant main effects of prosodic context, with the duration of the VOT lag being shorter at an IP-medial boundary than at an IP-initial one. The model also showed an effect of the interaction of duration of the stressed syllable and prosodic context: the longer the stressed syllable, the longer the VOT at an IP-medial boundary.



### 3.3 Vowel results

As mentioned above, variables pertaining to unstressed vowels only showed significant differences between contexts in Spanish and Portuguese (see Statistical summaries for English vowel results). Of these, duration was significantly different between IP-initial and IP-medial conditions in both Spanish and Portuguese, whereas F0 was significant for Spanish alone. F1 values failed to reach significance in any language (see Statistical summaries for all results).

#### 3.3.1 Vowel-related results in Spanish

Vowel duration values were evaluated through mixed effects models, with speaker and word as random effects. A total of 413 observations related to unstressed vowels in the Spanish data were analyzed. The best model for Spanish vowel durations (log-transformed) as the response variable included the following variables as fixed effects: duration of silent interval, word frequency, vowel height (two levels: high or low), and the interaction of duration of the stressed syllable and prosodic context (two levels: IP-initial or IP-medial). The `mixed()` function in R automatically calculated  $p$ -values. The final fixed and random effects estimates appear in Tables 7 and 8, respectively.

Table 7: Main effects for the Spanish model with vowel duration in target syllable as the response variable (log-transformed). The reference level for vowel height is “high”. Significance levels are \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$ ; n.s.: not significant.

Estimate	$\beta$	SE	$t$	$p(t)$
Intercept	1.388	0.018	10.89	<0.001***
Duration of pause	0.020	0.008	4.71	<0.001***
Vowel height = low	0.099	0.010	8.38	<0.001***
Stressed syllable duration x IP-initial	0.172	0.072	2.39	0.017*
Word frequency	-0.013	0.006	-2.22	0.01 *
Stressed syllable duration x IP-medial	-0.004	0.071	2.40	0.659 n.s.

As summarized in Table 7, results from the generalized linear model revealed significant main effects of vowel height, with the duration of the low vowel /a/ being overall longer than /i u/. The model also showed that a longer silent interval increases the duration of the vowel in the post-boundary CV syllable. Because silent pauses occurred most consistently before test words in the IP-initial

Table 8: Random effects intercept for Spanish /i a u/ duration (SD: standard deviation)

Variable	Variance	SD
Speaker	0.0088	0.0560
Word	0.0002	0.0050
Residual	0.0445	0.0853

condition, one can interpret the main effect of silent interval as an indirect correlation of domain boundary level with duration of unstressed vowels in the Spanish target syllables. There was also an effect of the interaction of duration of the stressed syllable and prosodic context: the longer the stressed syllable, the longer the vowel in the unstressed syllable at a prosodic boundary. Finally, the word frequency of the test word also showed the expected influence on duration: the higher the frequency, the shorter the unstressed vowel.

### 3.3.2 Vowel duration in Portuguese

In total, 400 tokens of Portuguese vowel duration data were fed into the mixed effects model. The model was fit through backward selection using all the applicable variables. Once again, predictors that failed to improve model fit (i.e.  $p \geq 0.05$ ) were excluded one at a time until the model described here was finalized. The best model for vowel duration (log-transformed) as the response variable in the Portuguese data included speaker and test item as random effects, and the following variables as fixed effects: duration of silent interval, vowel height (two levels: high or low), and the interaction of duration of the stressed syllable and prosodic context (two levels: IP-initial or IP-medial). The `mixed()` R function automatically calculated  $p$ -values. The final fixed and random effects estimates appear in Tables 9 and 10, respectively.

Results from the generalized linear model shown in Table 9 revealed significant main effects of vowel height, with the duration of the low vowel /a/ being overall longer than /i u/, similarly to the findings for Spanish. The model also showed that a longer pause duration is associated with increased duration of the vowel in the post-boundary CV syllable. There was also an effect of the interaction of duration of the stressed syllable and prosodic context: the longer the stressed syllable, the longer the vowel in the unstressed syllable at a prosodic boundary, more at an IP boundary than at the IP-medial domain.

Table 9: Main effects for the Portuguese model with vowel duration in the unstressed syllable as the response variable (log-transformed). The reference level for vowel height is “high”

Estimate	$\beta$	SE	$t$	$p(t)$
Intercept	2.464	0.419	17.88	<0.001***
Vowel height = low	1.591	0.109	10.70	<0.001***
Duration of pause	0.037	0.005	7.70	<0.001***
Stressed syllable duration x IP-initial	0.025	0.008	3.29	0.003**
Stressed syllable duration x IP-medial	0.018	0.010	1.83	0.298n.s.

Table 10: Random effects intercept – Duration of /i a u/ in target syllables (SD: standard deviation)

Variable	Variance	SD
Speaker	0.0018	0.0424
Word	0.0003	0.0164
Residual	0.0115	0.1071

## 4 Discussion

### 4.1 Implications for our understanding of prosodic structure

The current study is novel in two main ways: first, it provides evidence for the acoustic expression of domain-initial boundaries using a large sample. Secondly, by using the same materials to assess the phenomenon in three languages, it provides a direct evaluation of the claim that domain-initial boundaries manifest themselves in language-specific ways (e.g. Cho & McQueen 2005). Overall, this study corroborates the general hypothesis that words occurring immediately after a major prosodic boundary differ in their phonetic properties from those that follow a lower level boundary (Fougeron & Keating 1997). Furthermore, the different analyses presented above show that domain-initial strengthening operates on the phonetic properties of word-initial unstressed syllables, expanding them in language-specific ways.

As such, these results partially confirm two of the hypotheses that guided the current study (see Section 1.3). First, that domain-initial strengthening is not limited to articulation, and that acoustic variables are useful tools to describe (IP-) initial boundaries. Secondly, that different languages show particular correlates

of domain-initial strengthening. Within target CV syllables, English showed effects on the boundary-adjacent segment alone, whereas Spanish and Portuguese showed acoustic differences on both the consonant (i.e. occurrence of burst at stop release) and on the vowel (e.g. duration).

On the other hand, these language-specific characteristics manifested themselves differently from what had been predicted. Instead of differences in the magnitude of the effect, the current findings showed that both the type of segments affected (i.e. consonants in English, vowels in Spanish and Portuguese), and how those segments were affected (e.g. F0 in Spanish vowels, dispersion in Portuguese ones) differed. As such, these results reveal a somewhat inconsistent behavior of the variables under study. While these findings can be taken as confirmation of the phonological specificity of domain-initial strengthening, they render generalizations made over the entire sample much less straightforward. These and further limitations are taken up in more detail in Section 4.5 below.

The results presented here differ from those in previous literature in two relevant ways. First, domain-initial effects extended beyond the segment immediately following the boundary in Spanish and Portuguese, in which the vowel in the target CV syllable was longer in the IP-initial condition than IP-medially. Secondly, in the three languages, the stressed syllable, which was not boundary-adjacent, showed significant durational differences between prosodic contexts. Taken together, these results not only contradict the hypotheses laid out in Section 1.3, but also go against previous results suggesting that domain-initial strengthening effects are limited to the very first segment following the major boundary (Byrd et al. 2006, Cho & Keating 2009, Bombien et al. 2010, among others). As mentioned above, one key aspect of the current investigation is that it controlled for various levels of lexical stress and prosodic prominence. It is thus possible that the differences found here relate to the issue of prominence, taken up in more detail below.

## **4.2 Lexical stress and the locality of domain-initial strengthening**

This study was designed to isolate the influence of lexical stress and phrasal prominence from domain-initial strengthening. While this investigation focused on the segments in the boundary-adjacent unstressed syllable, the following stressed syllable in test words was measured as a control variable. The diagram in Figure 4 depicts the findings of the study for target unstressed syllables (represented as <cv>) while also showing the stressed syllable (i.e. <'CV>) in test words.

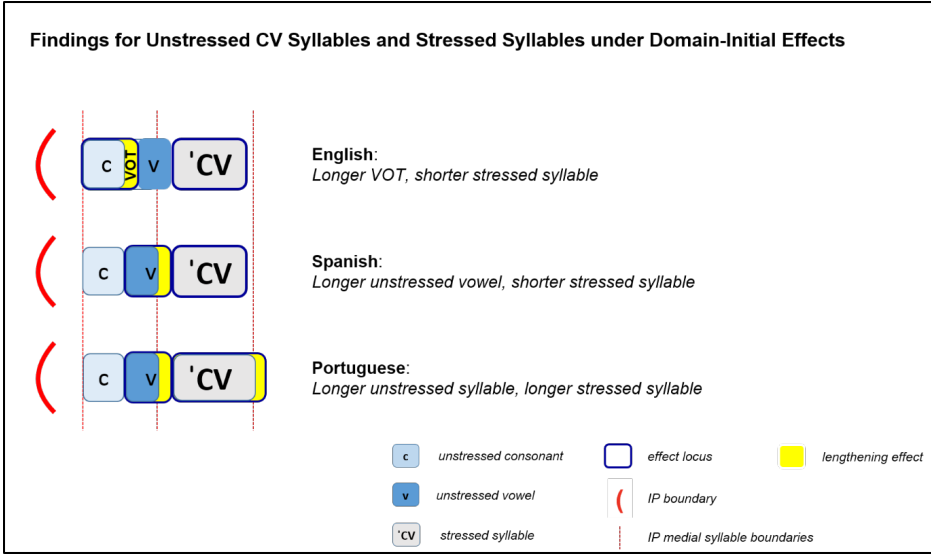


Figure 4: Schematic representation of the effects of boundary marking on IP-initial unstressed syllables, and the following stressed syllable for the languages in the study. Dashed lines represent syllable boundaries in the IP-medial condition as reference.

Individual statistical analyses for the three languages revealed that the stressed syllable had a significant main effect on the duration of VOT for English, and on vowel duration in both Spanish and Portuguese. Simply put, the effects observed on segments at/near the IP boundary were linked to the stressed syllable. While novel (to the best of my knowledge) in the domain-initial strengthening literature, these results have parallels in several studies of final lengthening that controlled for stress placement in *phrase-final* words.

The position of the lexically stressed syllable is a decisive factor in determining the scope of final lengthening in English (Kim et al. 2018, Cho et al. 2017, Byrd & Riggs 2008, Turk & Shattuck-Hufnagel 2007, White 2002, Oller 1973), German (Schubö & Zerbian 2023 [this volume]), Spanish (Rao 2010), and Portuguese (Frota 2000), but also in other stress languages such as Estonian (Krull 1997), Greek (Katsika 2016), Italian (Petrone et al. 2014), and Hebrew (Berkovits 1994). These studies and the current findings converge in that they all underscore the importance of lexical prominence in determining the scope of boundary-related effects. The present results are thus compatible with a view of prosodic boundaries and lexical prominence as closely related entities in the expression of prosodic structure (e.g. Turk & Shattuck-Hufnagel 2007, Katsika 2016).

Discussing domain-final boundaries, Turk & Shattuck-Hufnagel (2007) put forward a hypothesis that prosodic lengthening affects both the boundary-adjacent syllable and the stressed syllable that is not immediately adjacent to the boundary. According to this interpretation, the scope of boundary effects is determined by prosodic structure (i.e. the type of domain level) and by the phonological properties of the word (i.e. where stress is located) simultaneously. Put differently, the authors' hypothesis suggests that boundary marking is phonetically expressed with reference to lexical prominence. While articulatory in nature, Katsika's proposal (Katsika 2016) that prosodic boundaries and lexical prominence are integrated could also explain the data obtained in the current study. According to her, prosodic events related to boundary marking (i.e. domain-edge lengthening, articulatory strengthening, phrasal accents, boundary tones, and pauses) are interdependent, with lexical prosody functioning as the interface between phrasal prosody and constriction gestures (Katsika 2016: 169). Katsika's hypothesis is compelling as it can account not only for the acoustic results in the current study, but also for data in studies of domain-final lengthening and phrasal accent.

One possible interpretation of the aforementioned proposals is that phrasal effects would only affect segments in relation to a lexically prominent unit. This approach would imply that prosodic structure is phonetically cued from lower levels up, for instance from the Syllable to the Intonation Phrase. This bottom-up view of the prosodic hierarchy would suggest that smaller domains provide the framework upon which the whole structure is built. In light of the great deal of variability in prosodic phrasing, it may be useful to consider an approach that is more based on the concrete - and perhaps more stable - prosodic properties of lower-level domains such as the Word. Although formulated to account for other types of prosodic phenomena, Turk & Shattuck-Hufnagel (2007) and Katsika (2016) proposals would explain the current findings of lengthening occurring both in the segments immediately following the IP boundary and in the stressed syllable.

Viewed this way, the results of the present study suggest that the lexically prominent syllable may serve as an anchoring point for boundary marking, perhaps in similar ways to how it encodes phrasal prominence. The idea that domain-initial strengthening and lexical stress are interdependent coheres with the existing body of literature showing an association between lexical stress and phrasal accent in terms of pitch movement. In this interpretation, domain-initial effects would begin in the stressed syllable, and move leftwards to the phrase-initial boundary. That is, the locus of domain-initial effects would be best described as the stressed syllable, and the scope of the effect would potentially include segments between that syllable and the major prosodic boundary. Differences in

how lexical stress behaves phonetically would then explain specificities found in the implementation of prosodic structure, as suggested by Cho (2016).

### 4.3 The linguistic function of boundary effects

The evidence gathered in this study suggests that there is more to domain-initial effects than a purely biomechanical motivation. If domain-initial effects derived only from the start-up of articulation after a prosodic break, for instance, different languages would show consistent similarities in the way the prosodic effect operates on given segments such as /p t k/. The current results, as well as the findings from multiple studies reviewed in the Introduction, indicate that that is likely not the case. In other words, domain-initial effects differ in relevant ways from the marking of phrase edges before a prosodic boundary.

Domain-finally, the slowing down of articulators towards the end of the phrase suggests a physiological motivation behind pre-boundary effects such as phrase-final lengthening, or phrase-final creak. These phonetic effects can be interpreted as a reflection of the speaker's planning for the upcoming prosodic break, when most articulators will be at rest. This biomechanical process could then explain why phrase-final lengthening and/or phrase-final creak are crosslinguistically common (cf. Jun 2005, 2014, see also references in Garellek 2015).

On the other hand, the observation that the locus of phrase-final lengthening may relate to a lexically prominent syllable introduces a linguistic foundation for the effect. It is noteworthy that pitch movements that encode phrase-final edges also tend to associate with a linguistically relevant unit, such as a lexically prominent syllable in languages with lexical stress. As mentioned in the above discussion, the results of this investigation provide indication that prosodic boundary marking in English, Spanish, and Portuguese also relates to the lexical stress systems of these languages. The correlation of pre-boundary marking with word prosody thus suggests a possible parallel between the phonetic encoding of both edges of a prosodic domain.

The fact that the phonetic marking of the initial edge seems to relate to the segmental phonology of a language bespeaks a perhaps clearer linguistic motivation for domain-initial effects. Given the relevance of prosodic boundaries in speech recognition (Carlson 2009), an increase in phonemic contrast between neighboring segments at a phrase edge could possibly facilitate the parsing of speech. Because stressed syllables are prominent, it could be argued that they serve as a natural anchoring point for the marking of phrase edges – initial and final alike.

#### 4.4 Implications for phonological change: prestressed vowels in Portuguese

The data obtained here suggest that unstressed vowels are longer near an IP boundary in Portuguese, and that this boundary-related lengthening may prevent unstressed vowel reduction from taking place. These combined results may be useful to explain the asymmetry between prestressed and poststressed vowels in the language. As mentioned in the Introduction, Brazilian Portuguese shows a complex system of vowel qualities that relates simultaneously to lexical stress and syllable position within the word. Figure 5 illustrates the current variation in the expression of vowels in Brazilian Portuguese.

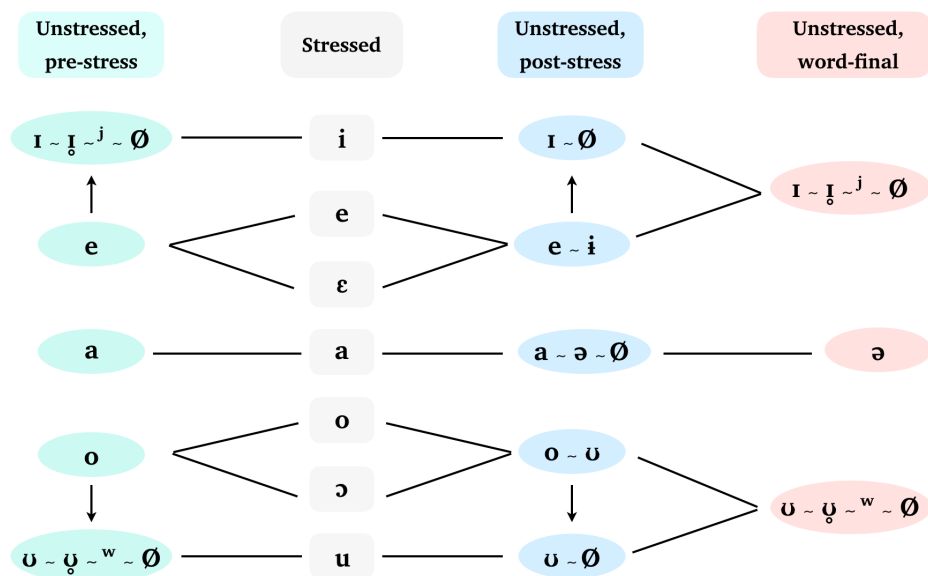


Figure 5: Variation in oral vowels in Brazilian Portuguese given stress placement within the word. Different colors represent variation in unstressed syllables with regards to word boundaries. Common allophonic variation shown within each oval; pointed arrows represent the direction of on-going sound change.

As can be seen in Figure 5, there is a stark contrast between unstressed oral vowels occurring in word-initial as opposed to word-final position. Up to five oral vowels can occur in unstressed word-initial syllables (i.e. /i e a o u/), whereas only /ɪ ə ʊ/ occur in unstressed final syllables. If duration is taken to be one of the most important factors in the neutralization of contrasts in vowels, the fact



that more vowel qualities are found in prestressed than poststressed position may be a consequence of domain-initial strengthening: the longer duration that can occur in prestressed position facilitates the distinction among more vowel qualities. Although the results regarding vowel dispersion were not significant for the current set of data, there is indication of a trend towards more dispersion at an IP-boundary than IP-medially, depicted in Figure 6.

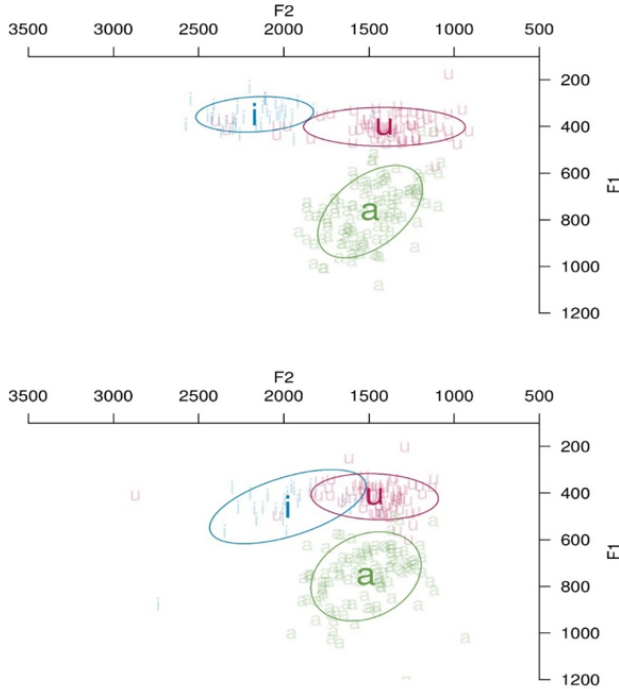


Figure 6: Vowel dispersion in Brazilian Portuguese /i a u/ given adjacency to an IP-boundary in the current study. The top graph shows vowels near the IP-initial boundary; bottom graph shows distribution of vowels IP-medially.

Put differently, the strengthening effect that derives from prosodic structure may be influencing patterns of lexical prominence in Portuguese (cf. Barnes 2006, Scheer & Ségéral 2008 for discussions). As the data (illustrated in Figure 6) suggests, domain-initial strengthening may be partially connected to the still relatively moderate reduction in prestressed syllables compared to poststressed ones in the language. Although more studies are necessary to evaluate this hypothesis, the current findings suggest that prosodic structure is at least a variable that must be taken into consideration in investigations of unstressed vowel reduction in Brazilian Portuguese, and in other languages.

#### 4.5 Limitations and future research

This study focused on the acoustic properties of voiceless stops and vowels immediately following prosodic domain boundaries, compared to sounds that are not adjacent to major boundaries. While some of the results discussed above may appear to contradict earlier articulatory studies, these findings are strictly limited to the acoustic realm, and no claims are made as to the locus, scope or anchor of domain-initial effects in terms of articulation. Future articulatory studies that manipulate lexical stress and accent using similar sample sizes would be useful to reconcile the findings put forward here with previous research on articulation.

Furthermore, more studies are needed to strengthen the present results for consonants other than voiceless stops. Similarly, future investigations of English and Portuguese should also test different vowel qualities in the prestressed CV syllable from the ones used here. For English, the fact that the test syllables always contained /ə/, a vowel which may not always show temporal expansion (see Cambier-Langeveld 1997 for analogous results concerning final lengthening in Dutch /ə/), may have influenced the scope of the boundary effect. For Portuguese, investigating whether proximity to the IP boundary affects the reduction of prestressed /e/ and /o/ to [i ɪ<sup>h</sup>] and [u ʊ<sup>w</sup>] would also help determine the extent to which boundary marking has implications to lexical stress reductions.

Since the current project focused on word-initial unstressed syllables, there was much less experimental control for the stressed syllable in the trisyllabic words used in the reading task. A follow-up investigation with tighter experimental control on both the unstressed and stressed syllables could potentially increase the validity of the present findings. As explained above, this study was not explicitly designed to capture differences between stressed syllables, but main effects on stressed syllables were found in all of the languages investigated. More research is hence necessary to confirm the associations found between prosodic context and the phonetic characteristics of stressed syllables not immediately following a major prosodic boundary. A future study could also contrast unstressed and stressed CV syllables like the ones tested here.

Acoustic investigations that manipulate the number of unstressed syllables between the boundary and the stressed syllable would also offer important contributions to our understanding of domain-initial strengthening effects. Additionally, although the test words measured in this research did not receive the main phrasal accent, no specific control was undertaken with regards to the presence and type of prenuclear pitch accents. Future studies that manipulate pitch accent type and placement would constitute a relevant refinement of the methods employed here, including regarding the use of focus statements in both experimental conditions.

Finally, this investigation only considers data from English, Spanish, and Portuguese, all of which are languages with somewhat unpredictable lexical stress. It is possible that a similar acoustic study of languages with other types of word prosody systems, including languages with fixed lexical stress, may result in different associations between boundary marking and lexical prominence. What's more, the claims made here may not be applicable to varieties of these languages other than the ones investigated here, namely American English, Mexican Spanish, and Brazilian Portuguese, given the known prosodic differences between dialects of the same language (e.g. Clopper & Smiljanic 2011: 145, for English; Prieto & Roseano 2010, for Spanish; Frota et al. 2015 for Portuguese). In the case of Spanish, it would be desirable to conduct a similar study using only monolingual speakers instead of the fully bilingual participants recorded for the current project.

## 5 Conclusion

This study sought to shed new light on the acoustic manifestation of domain-initial strengthening, a type of boundary-induced prominence, from a crosslinguistic perspective. In doing so, it sought to provide more data on the ways through which prosodic structure organizes speech. In contrast to most previous investigations on the topic, the current project looked at unstressed segments occurring in words that did not bear the main (nuclear) phrasal accent. The data obtained from 42 speakers of English, Spanish, and Portuguese revealed that the acoustic correlates of boundary marking extend beyond the initial segment in unstressed CV syllables, affecting the vowel in Spanish and Portuguese, and the stressed syllable in all three languages.

The data discussed here suggest a close connection between the grouping and prominence functions of prosody, in which the stressed syllable may serve as the anchoring point for boundary marking. This proposal is in line with findings from studies of domain-final effects that controlled for stress placement in test words (Turk & Shattuck-Hufnagel 2007, Cho et al. 2013, Katsika 2016). These investigations show that phrase-final lengthening is initiated farther away from the boundary in polysyllabic words that do not have stress on the final syllable. The combined evidence seems to suggest that lexically stressed syllables play a role in marking both domain-initial and domain-final boundaries of major phrases. This function of the stressed syllable would then add to its already established function in marking phrase-level prominence in some languages, and thus provide support to the view that prosodic structure manifests itself phonetically through the interaction of segmental and suprasegmental factors.

More broadly, the results of the current study reinforce the idea that speakers actively use language-specific phonological knowledge (e.g. VOT lag for English /p t k/, vowel duration for Spanish and Portuguese /i a u/) to implement phonetic distinctions that are relevant to speech categories (e.g. Kingston & Diehl 1994, Cho & Ladefoged 1999). The present findings corroborate the hypothesis that speakers indicate the grouping of their speech units by manipulating phonetic detail (Cho 2016), thereby highlighting the effects of prosody on segmental phonetics (i.e. the prosody-phonetics interface). Finally, this investigation presents further evidence that phonetic information relates to multiple levels of prosodic structure simultaneously.

## Acknowledgements

I would like to thank Caroline L. Smith, and Taehong Cho for their insightful comments on an early version of this manuscript. I am also grateful to the anonymous reviewers for their suggestions and constructive criticism, all of which helped improve the chapter.

## Appendix A English stimuli

### A.1 /p/

#### (1) *petitions*

- a. You're talking about ordinary polls but it's not the same || petitions must be SIGNED to be valid
- b. Very often students do get to voice their concerns but these silly # petitions make no difference

#### (2) *patrolmen*

- a. Policemen can certainly arrest you but that's not the case || patrolmen only REINFORCE order
- b. I'm used to being stopped by the police but I have to say: those angry # patrolmen really scared me

### A.2 /t/

#### (3) *tequila*

- a. I've checked labels plus I've tried both drinks, so I'm pretty certain || tequila is WEAKER than pure vodka

- b. I'm used to drinking strong liquor because I don't like beer but that nasty # tequila made me so sick
- (4) *toccata(s)*
  - a. No, they're not something you eat at all! I know from music class || toccatas are just long MUSIC pieces
  - b. I did enjoy it; she's an excellent musician, no doubt: that classy # toccata was fantastic

### A.3 /k/

- (5) *katrina*
  - a. A lot of hurricanes do hit those parts but this time you are wrong || Katrina hit NEW ORLEANS, not Texas
  - b. I've lived through many horrible storms that caused much damage but that deadly # Katrina destroyed the land
- (6) *capacious*
  - a. It doesn't refer to ability! You can check for yourself || capacious means ROOMY or full of space
  - b. It is very sad there's not too much they can do at this point: the city's # capacious museum closed

## Appendix B Spanish stimuli

### B.1 /p/

- (7) *pitada* [pi'ta.ða]
  - a. Estás confundido || pitada quiere decir SOPLADO más que sonido o pitido  
'You're mistaken, whistling is a BLOWING SOUND more than a noise or a beep'
  - b. A causa de la lluvia, el árbitro Federico dio la # pitada a las tres horas  
'Because of the rain Federico the referee blew the whistle to end the match at 3 o'clock'
- (8) *patrulla* [pa'truja]
  - a. Aquí en México || patrulla quiere decir un CARRO de vigilancia en la ciudad  
'Here in Mexico, a patrol is a CAR used by city police'

- b. A pesar de las protestas, el gobierno va a mantener la # patrulla policial diaria  
'Despite the demonstrations, the government is keeping the daily police patrols'

**B.2 /t/**

- (9) *tipazo* [ti'pa.so]
  - a. No te confundas || tipazo quiere decir AMABLE más que un cuerpo atractivo  
'Don't mix the two up, a stud is more like a NICE guy than a hot one'
  - b. Es un tanto vulgar, aquí en esta zona no se dice # tipazo a las personas  
'That's a little vulgar; around here we don't call anyone a stud'
- (10) *tacada* [ta'ka.ða]
  - a. Según sus abuelos || tacada tiene que ver con ARMAS de fuego y no con el billar  
'According to his grandparents, a strike is something to do with GUNS, not with playing pool'
  - b. Ganó el partido porque su papá le enseñó una # tacada espectacular  
'S/he won the match because her/his dad taught her/him a great move'

**B.3 /k/**

- (11) *cuchara* [ku'tʃara]
  - a. Aprendí con ellos || cuchara se refiere TAMBIÉN a la herramienta del albañil  
'I learned this from them "cuchara" ALSO means a trowel that you use to build stuff'
  - b. Los albañiles estuvieron varias horas buscando la # cuchara para el muro  
'The contractors spent several hours looking for a trowel to build the wall'
- (12) *capricho* [ka'pɾitʃo]
  - a. Me parece raro || capricho significa un DESEO irracional muy intenso  
'That sounds strange, a whim means an irrational DESIRE that is very intense'

- b. Los abuelos prepararon recetas para cumplirle su # capricho gastronómico  
'Her/His grandparents cook recipes just to satisfy her/his food whims'

## Appendix C Portuguese stimuli

### C.1 /p/

- (13) *pitada* [pɪ'ta.də]  
a. Tem muito sal aqui || pitada quer dizer só UM POUQUINHO do ingrediente na receita  
'You put too much salt in this; a pinch means JUST A LITTLE of the ingredient'  
b. Pimenta caiena é mais forte do que do reino; só uma # pitada tá mais que bom  
'Cayenne pepper is much stronger than black pepper; just a pinch is more than enough'
- (14) *patola* [pa'tɔ.lə]  
a. Não é um pato não || patola tem a ver com TAMANHO ou peso duma pessoa  
'It doesn't mean full of stock, stocky has to do with someone's SIZE or weight'  
b. Elas venderam todos os filhotes mas essa cachorrinha # patola ninguém levou  
'They sold most of the puppies but no one really wanted to take the stocky one'

### C.2 /t/

- (15) *tutela* [tʊ'tɛ.lə]  
a. Isso é outra coisa || tutela garante a AUTORIDADE sobre uma criança  
'That's something else entirely; guardianship means having LEGAL AUTHORITY over a child'  
b. Meu pai ficou sabendo outro dia que o Gilberto perdeu a # tutela dos três filhos  
'My father heard the other day that Gilberto lost custody of his three children'

(16) *tacada* [ta'ka.də]

- a. Esquece de taco || tacada quer dizer uma IDEIA inteligente que deu certo  
'Forget about the word taco; tacada means a clever IDEA that panned out'
- b. Mesmo sem conhecer o gerente, não dá pra negar que aquela # tacada foi de mestre  
'You don't have to know the manager to acknowledge that his clever move was exceptional'

C.3 /k/

(17) *cutelo* [kʊ'tɛ.lɐ]

- a. Não é de açougue || cutelo é meio que um facão PEQUENO de uso diário  
'It's not a butcher knife, a cleaver is a kind of SMALL hatchet for daily use in the kitchen'
- b. Dependendo do tipo de carne é melhor usar aquele # cutelo maiorzinho  
'I guess it depends on the kind of meat but you should probably use that largish cleaver over there'

(18) *capela* [ka'pɛ.lə]

- a. Não é igrejinha || capela é um nicho PEQUENO dedicado a algum santo  
'It's not a small church, a chapel is a small area dedicated to a given Catholic saint'
- b. De todas as partes da igreja a que eu mais gosto é aquela # capela dourada lá  
'Of all the areas of the church, my favorite spot is that golden chapel over there'



## Appendix D Statistical summaries

Table 11: Statistics summary for the English data (ms: milliseconds; dur: duration;  $\tilde{m}$ : median; MAD: median absolute deviation).

Variables	Prosodic context							
	IP-initial				IP-medial			
	min	max	$\tilde{m}$	MAD <sup>a</sup>	min	max	$\tilde{m}$	MAD
VOT /p t k/ (ms)	8	84	36	12	7	83	32	13
Vowel dur (ms)	15	73	32	12	21	71	33	12
Vowel F1 (normalized) <sup>b</sup>	7.73	22.4	9.95	1.33	6.23	26.54	9.96	1.81
Dur pause (ms)	32	1187	101	148	0	48	1	0
Dur stressed syllable (ms)	102	329	201	40	131	346	210	39

<sup>a</sup>MAD: median absolute deviation. MAD is a more robust measure of variability in non-normal distributions than the standard deviation (Levshina 2015).

<sup>b</sup>Obtained for each token based on the means and standard deviations calculated over all productions by the same speaker + 10.

Table 12: Statistics summary for the Spanish data (ms: milliseconds; dur: duration;  $\tilde{m}$ : median; MAD: median absolute deviation).

Variables	Prosodic context							
	IP-initial				IP-medial			
	min	max	$\tilde{m}$	MAD	min	max	$\tilde{m}$	MAD
VOT /p t k/ (ms)	4	62	15	6	4	46	17	7
Vowel dur (ms)	23	99	59	16	24	98	54	14
Vowel F1 (normalized)	8.14	13.87	9.86	1.36	8.36	13.15	9.56	1.16
Dur pause (ms)	59	1164	125	131	0	89	3	0
Dur stressed syllable (ms)	109	375	201	31	142	364	212	30

Table 13: Statistics summary for the Portuguese data (ms: milliseconds; dur: duration; MAD: median absolute deviation).

Variables	Prosodic context							
	IP-initial				IP-medial			
	min	max	ñ	MAD	min	max	ñ	MAD
VOT /p t k/ (ms)	7	58	20	12	1	51	22	13
Vowel dur (ms)	19	117	56	21	16	100	45	21
Vowel F1 (normalized)	8.07	12.27	9.95	1.02	8.20	12.96	10.21	1.26
Dur pause (ms)	47	724	185	152	0	33	0	0
Dur stressed syllable (ms)	120	468	246	71	133	385	209	37

## Appendix E Wilcoxon tests results

Table 14: Results of two-tailed Wilcoxon tests for statistically significant differences between experimental conditions (IP-initial vs. IP-medial position).

Variable	English		Spanish		Portuguese	
	<i>W</i>	<i>p</i>	<i>W</i>	<i>p</i>	<i>W</i>	<i>p</i>
VOT /p t k/	26472	0.015*	19958	0.265	22193	0.917
Vowel duration	27261	0.487	25630	<0.001***	30646	<0.001***
Vowel F1	22022	0.917	22846	0.178	20057	0.098
Dur pause	51859	<0.001***	40296	<0.001***	42210	<0.001***
Dur stressed syllable	23184	<0.001***	18090	<0.001***	30630	<0.001***

## References

- Abramson, Arthur S. & Douglas H. Whalen. 2017. Voice onset time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics* 63. 75–86. DOI: 10.1016/j.wocn.2017.05.002.
- Ahn, Suzy. 2018. The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazilian Portuguese. *Journal of Phonetics* 71. 451–467. DOI: 10.1016/j.wocn.2018.10.003.

- Alves, Mariane Antero, Izabel Christine Seara, Fernando Santana Pacheco, Simone Klein & Rui Seara. 2008. On the voiceless aspirated stops in Brazilian Portuguese. In *Computational processing of the Portuguese language (PROPOR 2008)*, 248–251. Springer. DOI: 10.1007/978-3-540-85980-2\_31.
- Audacity Team. 2014. *Audacity: Free audio editor and recorder*. <https://www.audacityteam.org/>.
- Avelino, Heriberto. 2018. Mexico City Spanish. *Journal of the International Phonetic Association* 48(2). 223–230. DOI: 10.1017/S0025100316000232.
- Barnes, Jonathan A. 2006. *Strength and weakness at the interface: Positional neutralization in phonetics and phonology*. Berlin/NewYork: Mouton de Gruyter. DOI: 10.1515/9783110197617.
- Beckman, Mary E. 1986. *Stress and non-stress accent*. Dordrecht: Foris Publications. DOI: 10.1515/9783110874020.
- Beckman, Mary E., Kenneth J. de Jong, Sun-Ah Jun & Sook-Hyang Lee. 1992. The interaction of coarticulation and prosody in sound change. *Language and Speech* 35(1). 45–58. DOI: 10.1177/002383099203500205.
- Beckman, Mary E. & Jan Edwards. 1994. Articulatory evidence for differentiating stress categories. In *Papers in Laboratory Phonology III: Phonological structure and phonetic form*, 7–33. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511659461.002.
- Beckman, Mary E. & Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309. DOI: 10.1017/s095267570000066x.
- Bell, Alan, Jason M. Brenier, Michelle Gregory, Cynthia Girand & Dan Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60(1). 92–111. DOI: 10.1016/j.jml.2008.06.003.
- Beňuš, Štefan & Juraj Šimko. 2014. Emergence of prosodic boundary: Continuous effects of temporal affordance on inter-gestural timing. *Journal of Phonetics* 44. 110–129. DOI: 10.1016/j.wocn.2013.12.005.
- Berkovits, Rochele. 1994. Durational effects in final lengthening, gapping and contrastive stress. *Language and Speech* 37. 237–250. DOI: 10.1177/002383099403700302.
- Bodur, Kübra, Sweeney Branje-Canioni, Morgane Peirollo, Ingrid Tiscareno & James German. 2021. Domain-initial strengthening in Turkish: Acoustic cues to prosodic hierarchy in stop consonants. In *Proceedings of InterSpeech*, 1459–1463. DOI: 10.21437/interspeech.2021-2230.
- Boersma, Paul & David Weenink. 2019. *Praat: Doing phonetics by computer*. <http://www.praat.org/>.

- Bombien, Lasse, Christine Mooshammer, Phil Hoole & Barbara Kühnert. 2010. Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics* 38(3). 388–403. DOI: 10.1016/j.wocn.2010.03.003.
- Byrd, Dani, Jelena Krivokapić & Sungbok Lee. 2006. How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America* 120. 1589–1599. DOI: 10.1121/1.2217135.
- Byrd, Dani & Daylen Riggs. 2008. Locality interactions with prominence in determining the scope of phrasal lengthening. *Journal of the International Phonetic Association* 38(2). 187–202. DOI: 10.1017/s0025100308003460.
- Câmara JR, Joaquim Mattoso. 1972. *The Portuguese language*. Chicago: University of Chicago Press.
- Cambier-Langeveld, Tina. 1997. The domain of final lengthening in the production of Dutch. In Jane A. Coerts & Helen de Hoop (eds.), *Linguistics in the Netherlands*, 13–24. Amsterdam: Benjamins. DOI: 10.1075/avt.14.04cam.
- Cantoni, Maria Mendes. 2013. *O acento no português brasileiro: uma abordagem experimental [Accent in Brazilian Portuguese: An experimental approach]*. Belo Horizonte, Brazil: Universidade Federal de Minas Gerais. (Doctoral dissertation). <http://hdl.handle.net/1843/LETR-96NLVH>.
- Carlson, Katy. 2009. How prosody influences sentence comprehension. *Language and Linguistics Compass* 3(5). 1188–1200. DOI: 10.1111/j.1749-818x.2009.00150.x.
- Cho, Taehong. 2011. Laboratory phonology. In Nancy C. Kula, Bert Botma & Kuniya Nasukawa (eds.), *The Continuum Companion to phonology*, 343–368. London: Continuum.
- Cho, Taehong. 2015. Language effects on timing at the segmental and suprasegmental levels. In Melissa A. Redford (ed.), *The handbook of speech production*, 505–559. Chichester: Wiley-Blackwell. DOI: 10.1002/9781118584156.ch22.
- Cho, Taehong. 2016. Prosodic boundary strengthening in the phonetics–prosody interface. *Language and Linguistics Compass* 10(3). 120–141. DOI: 10.1111/lnc3.12178.
- Cho, Taehong & Patricia A. Keating. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29(2). 155–190. DOI: 10.1006/jpho.2001.0131.
- Cho, Taehong & Patricia A. Keating. 2009. Effects of initial position versus prominence in English. *Journal of Phonetics* 37(4). 466–485. DOI: 10.1016/j.wocn.2009.08.001.
- Cho, Taehong, Daejin Kim & Sahyang Kim. 2017. Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics* 64. 71–89. DOI: 10.1016/j.wocn.2016.12.003.

- Cho, Taehong, Jiseung Kim & Sahyang Kim. 2013. Preboundary lengthening and preaccentual shortening across syllables in a trisyllabic word in English. *The Journal of the Acoustical Society of America* 133. 384–390. DOI: 10.1121/1.4800179.
- Cho, Taehong & Peter Ladefoged. 1999. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27. 207–229. DOI: 10.1006/jpho.1999.0094.
- Cho, Taehong & James M. McQueen. 2005. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics* 33(2). 121–157. DOI: 10.1016/j.wocn.2005.01.001.
- Cho, Taehong, James M. McQueen & Ethan A. Cox. 2007. Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35(2). 210–243. DOI: 10.1016/j.wocn.2006.03.003.
- Cho, Taehong, Douglas H. Whalen & Gerard Docherty. 2019. Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of Phonetics* 72. 52–65. DOI: 10.1016/j.wocn.2018.11.002.
- Clopper, Cynthia G. & Rajka Smiljanic. 2011. Effects of gender and regional dialect on prosodic patterns in American English. *Journal of Phonetics* 39(2). 160. DOI: 10.1016/j.wocn.2011.02.006.
- Cole, Jennifer. 2015. Prosody in context: A review. *Language, Cognition and Neuroscience* 30(1-2). 1–31. DOI: 10.1080/23273798.2014.963130.
- Cole, Jennifer, Heejin Kim, Hansook Choi & Mark Hasegawa-Johnson. 2007. Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics* 35(2). 180–209. DOI: 10.1016/j.wocn.2006.03.004.
- Cristófar-Silva, Thaïs, Izabel Christine Seara, Adelaide Silva, Andreia Schurt Rauber & Maria Cantoni. 2019. *Fonética acústica: Os sons do Português brasileiro [Acoustic phonetics: The sounds of Brazilian Portuguese]*. São Paulo: Contexto.
- Crosswhite, Katherine M. 2001. *Vowel reduction in Optimality Theory*. New York: Routledge.
- Crystal, Thomas H. & Arthur S. House. 1988. The duration of American-English vowels: An overview. *Journal of Phonetics* 16(3). 263–284. DOI: 10.1016/s0095-4470(19)30500-5.
- Davidson, Lisa. 2006. Schwa elision in fast speech: Segmental deletion or gestural overlap? *Phonetica* 63(2-3). 79–112. DOI: 10.1159/000095304.
- Davies, Mark. 2008. The corpus of contemporary American English (COCA): 560 million words. <https://www.english-corpora.org/coca/>.

- Davies, Mark. 2016. Corpus del Espanol: Web/dialects. <https://www.corpusdelespanol.org>.
- Davies, Mark & Michael Ferreira. 2006. Corpus do Portugues. <https://www.corpusdoportugues.org>.
- de la Mota, Carme, Pedro M. Butragueño & Pilar Prieto. 2010. Mexican Spanish intonation. In Pilar Prieto & Paolo Roseano (eds.), *Transcription of intonation of the Spanish language*, 319–350. Munich: LINCOM Europa. [http://prosodia.upf.edu/home/arxiu/activitats/4th\\_workshop/protegit/Mexican\\_final.pdf](http://prosodia.upf.edu/home/arxiu/activitats/4th_workshop/protegit/Mexican_final.pdf).
- de Jong, Kenneth J. 2004. Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics* 32(4). 493–516. DOI: 10.1016/j.wocn.2004.05.002.
- de Jong, Nivja H. & Ton Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41(2). 385–390. DOI: 10.3758/brm.41.2.385.
- Delforge, Ann Marie. 2008. Gestural alignment constraints and unstressed vowel devoicing in Andean Spanish. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 147–155. <http://www.lingref.com/cpp/wccfl/26/paper1668.pdf>.
- Dilley, Laura, Stefanie Shattuck-Hufnagel & Mari Ostendorf. 1996. Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* 24(4). 423–444. DOI: 10.1006/jpho.1996.0023.
- Erickson, Donna. 1998. Effects of contrastive emphasis on jaw opening. *Phonetica* 55(3). 147–169. DOI: 10.1159/000028429.
- Fletcher, Janet. 2010. The prosody of speech: Timing and rhythm. In William J. Hardcastle, John Laver & Fiona E. Gibbon (eds.), *The handbook of phonetic sciences*, 521–602. Chichester: Wiley. DOI: 10.1002/9781444317251.ch15.
- Fónagy, Ivan & Klara Magdics. 1960. Speed of utterance in phrases of different lengths. *Language and Speech* 3(4). 179–192. DOI: 10.1177/002383096000300401.
- Fougeron, Cecile. 1999. Prosodically conditioned articulatory variations: A review. *UCLA Working Papers in Phonetics* 97. 1–73. <https://escholarship.org/uc/item/3pg0w66x>.
- Fougeron, Cecile. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29(2). 109–135. DOI: 10.1006/jpho.2000.0114.
- Fougeron, Cecile & Patricia A. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America* 101(6). 3728–3740. DOI: 10.1121/1.418332.

- Fowler, Carol A. & Jonathan Housum. 1987. Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26(5). 489–504. DOI: 10.1016/0749-596X(87)90136-7.
- Frota, Sonia. 2000. *Prosody and focus in European Portuguese: Phonological phrasing and intonation*. New York/London: Garland Pub. DOI: 10.4324/9781315054384.
- Frota, Sonia, Marisa Cruz, Flaviane Fernandes-Svartman, Gisela Collischonn, Aine Fonseca, Carolina Serra, Pedro Oliveira & Marina Vigário. 2015. Intonational variation in Portuguese: European and Brazilian varieties. In Sonia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 235–283. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199685332.003.0007.
- Frota, Sonia, Mariapola D'Imperio, Gorka Elordieta, Pilar Prieto & Marina Vigário. 2007. The phonetics and phonology of intonational phrasing in Romance. In *Segmental and prosodic issues in Romance phonology*, 131–152. Benjamins. DOI: 10.1075/cilt.282.10fro.
- Frota, Sonia & Joao Moraes. 2016. Intonation of European and Brazilian Portuguese. In W. Leo Wetzels, Joao Costa & Sergio Menuzzi (eds.), *The handbook of Portuguese linguistics*, 141–166. Cambridge, MA: Wiley-Blackwell. DOI: 10.1002/9781118791844.ch9.
- Frota, Sonia & Pilar Prieto. 2015. *Intonation in Romance*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199685332.001.0001.
- Garellek, Marc. 2015. Perception of glottalization and phrase-final creak. *The Journal of the Acoustical Society of America* 137(2). 822–831. DOI: 10.1121/1.4906155.
- Georgeton, Laurianne & Cecile Fougeron. 2014. Domain-initial strengthening on French vowels and phonological contrasts: Evidence from lip articulation and spectral variation. *Journal of Phonetics* 44. 83–95. DOI: 10.1016/j.wocn.2014.02.006.
- Goldman, Jean-Philippe. 2011. EasyAlign: An automatic phonetic alignment tool under Praat. In *Proceedings of InterSpeech*. DOI: 10.21437/interpeech.2011-815.
- Gries, Stefan Thomas. 2013. *Statistics for linguistics with R: A practical introduction*. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110307474.
- Gussenhoven, Carlos & Antonius C. M. Rietveld. 1992. Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics* 20. 283–303. DOI: 10.1016/s0095-4470(19)30636-9.
- Hsu, Chai-Shune & Sun-Ah Jun. 1998. Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? *UCLA Working Papers in Phonetics* 96. 69–89.
- Hualde, José Ignacio. 2007. Stress removal and stress addition in Spanish. *Journal of Portuguese Linguistics* 5(6). 59–89. DOI: 10.5334/jpl.145.

- Hualde, José Ignacio. 2012. Stress and rhythm. In José Ignacio Hualde, Antxon Olarrea & Erin O'Rourke (eds.), *The handbook of Hispanic linguistics*, 153–171. Chichester: Wiley-Blackwell. DOI: 10.1002/9781118228098.ch8.
- Hualde, José Ignacio & Marianna Nadeu. 2014. Rhetorical stress in Spanish. In Harry van der Hulst (ed.), *Word stress: Theoretical and typological issues*, 228–252. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9781139600408.010.
- Hualde, José Ignacio & Pilar Prieto. 2015. Intonational variation in Spanish: European and American varieties. In Sonia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 350–392. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199685332.003.0010.
- Jun, Sun-Ah. 2005. *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199249633.001.0001.
- Jun, Sun-Ah. 2006. Intonation. In Lynn Nadel (ed.), *Encyclopedia of cognitive science*, 618–624. Chichester: Wiley-Blackwell. DOI: 10.1002/0470018860.s00241.
- Jun, Sun-Ah. 2014. *Prosodic typology II: The phonology of intonation and phrasing*. Sun-Ah Jun (ed.). Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199567300.001.0001.
- Jun, Sun-Ah & Mary E. Beckman. 1994. Distribution of devoiced high vowels in Korean. In *Proceedings of the 3rd International Conference on Spoken Language Processing*, 479–482. DOI: 10.21437/icslp.1994-110.
- Katsika, Argyro. 2016. The role of prominence in determining the scope of boundary-related lengthening in Greek. *Journal of Phonetics* 55. 149–181. DOI: 10.1016/j.wocn.2015.12.003.
- Keating, Patricia A. 2004. Phonetic encoding of prosodic structure. *UCLA Working Papers in Phonetics* 103. 48–63. <http://phonetics.linguistics.ucla.edu/workpapph/103/4-Keating-ISSPSydney.pdf>.
- Keating, Patricia A., Taehong Cho, Cecile Fougeron & Chai-Shune Hsu. 2003. Domain-initial strengthening in four languages. In *Papers in Laboratory Phonology VI: Phonetic interpretation*, 145–163. Cambridge: Cambridge University Press.
- Kessinger, Rachel H. & Sheila E. Blumstein. 1998. Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *Journal of Phonetics* 26(2). 117–128. DOI: 10.1006/jpho.1997.0069.
- Kim, Sahyang, Jiseung Kim & Taehong Cho. 2018. Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English. *Journal of Phonetics* 71. 65–80. DOI: 10.1016/j.wocn.2018.07.004.



- Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70(3). 419–454. DOI: 10.2307/416481.
- Krivokapić, Jelena. 2014. Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1658). 1–10. DOI: 10.1098/rstb.2013.0397.
- Krull, Diana. 1997. Prepausal lengthening in Estonian: Evidence from conversational speech. In Ilse Lehiste & Jaan Ross (eds.), *Estonian prosody: Papers from a symposium*, 136–148.
- Kuzla, Claudia, Taehong Cho & Mirjam Ernestus. 2007. Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics* 35(3). 301–320. DOI: 10.1016/j.wocn.2006.11.001.
- Ladd, D. Robert. 2008. *Intonational phonology*. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9780511808814.
- Lavoie, Lisa M. 2001. *Consonant strength: Phonological patterns and phonetic manifestations*. New York: Garland. 254. DOI: 10.4324/9780203826423.
- Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: Benjamins. DOI: 10.1075/z.195.
- Liberman, Mark Y. & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8(2). 249–336. <https://www.jstor.org/stable/4177987>.
- Lisker, Leigh & Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *WORD* 20(3). 384–422. DOI: 10.1080/00437956.1964.11659830.
- Mateus, Maria Helena & Ernesto D’Andrade. 2000. *The phonology of Portuguese*. Oxford: Oxford University Press.
- Mo, Yoonsook. 2007. Temporal, spectral evidence of devoiced vowels in Korean. In *Proceedings of the 16th ICPHS*, 445–448. <http://www.icphs2007.de/conference/Papers/1597/1597.pdf>.
- Nadeu, Marianna. 2014. Stress- and speech rate-induced vowel quality variation in Catalan and Spanish. *Journal of Phonetics* 46. 1–22. DOI: 10.1016/j.wocn.2014.05.003.
- Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris Publications. DOI: 10.1515/9783110977790.
- Oh, Sejin. 2021. Phonetic and phonological vowel reduction in Brazilian Portuguese. *Phonetica*. 435–465. DOI: 10.1515/phon-2021-2012.
- Oller, D. Kimbrough. 1973. The effect of position in utterance on speech segment duration in English. *The Journal of the Acoustical Society of America* 54(5). 1235–1247. DOI: 10.1121/1.1914393.

- Onaka, Akiko, Sallyanne Palethorpe, Catherine Watson & Jonathan Harrington. 2003. An acoustic analysis of domain-initial strengthening effect in Japanese. In *Proceedings of the sixth international seminar on speech production*, 201–206.
- Ortega-Llebaria, Marta & Pilar Prieto. 2007. Disentangling stress from accent in Spanish: Production patterns of the stress contrast in deaccented syllables. In *Segmental and prosodic issues in Romance phonology*, 155–175. Amsterdam: Benjamins. DOI: 10.1075/cilt.282.11ort.
- Parrell, Benjamin. 2014. *Dynamics of consonant reduction*. University of Southern California. (Doctoral dissertation).
- Petrone, Caterina, Mariapola D’Imperio, Leonardo Lancia & Susanne Fuchs. 2014. The interplay between prosodic phrasing and accentual prominence on articulatory lengthening in Italian. In *Proceedings of Speech Prosody*, 192–196. DOI: 10.21437/speechprosody.2014-26.
- Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. Massachusetts Institute of Technology. (Doctoral dissertation). <http://dspace.mit.edu/handle/1721.1/16065>.
- Pierrehumbert, Janet B. & Julia Bell Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 271–312. Cambridge, MA: MIT Press. DOI: 10.7916/D8KD24FP.
- Pierrehumbert, Janet B. & David Talkin. 1992. Lenition of /h/ and glottal stop. In *Papers in Laboratory Phonology II: Gesture, segment, prosody*, 90–117. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9780511519918.005.
- Plag, Ingo, Gero Kunter & Mareile Schramm. 2011. Acoustic correlates of primary and secondary stress in North American English. *Journal of Phonetics* 39(3). 362–374. DOI: 10.1016/j.wocn.2011.03.004.
- Prieto, Pilar & Paolo Roseano. 2010. *Transcription of intonation of the Spanish language*. Munich: LINCOM Europa.
- Prieto, Pilar, Jan van Santen & Julia Hirschberg. 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23(4). 429–51. DOI: 10.1006/jpho.1995.0032.
- R Core Team. 2018. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Ramanarayanan, Vikram, Erik Bresch, Dani Byrd, Louis Goldstein & Shrikanth S. Narayanan. 2009. Analysis of pausing behavior in spontaneous speech using real-time magneticresonance imaging of articulation. *The Journal of the Acoustical Society of America* 126(5). EL160–EL165. DOI: 10.1121/1.3213452.
- Rao, Rajiv. 2010. Final lengthening and pause duration in three dialects of Spanish. In *Selected proceedings of the 4th conference on laboratory approaches to Spanish phonology*, 69–82. <http://www.lingref.com/cpp/lasp/4/paper2368.pdf>.

- Redi, Laura & Stefanie Shattuck-Hufnagel. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29(4). 407–429. DOI: 10.1006/jpho.2001.0145.
- Roettger, Timo & Matthew Gordon. 2017. Methodological issues in the study of word stress correlates. *Linguistics Vanguard* 3. 1. DOI: 10.1515/lingvan-2017-0006.
- Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, Scott Seyfarth, Kyle Gorman, Hilary Prichard & Jiahong Yuan. 2014. *FAVE: Forced alignment and vowel extraction*). Version 1.1.3. DOI: 10.5281/zenodo.9846.
- Scheer, Tobias & Philippe Ségéral. 2008. Positional factors in lenition and fortition. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and fortition*, 131–172. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110211443.1.131.
- Schubö, Fabian & Sabine Zerbian. 2023. The patterns of pre-boundary lengthening in German. In Fabian Schubö, Sabine Zerbian, Sandra Hanne & Isabell Wartenburger (eds.), *Prosodic boundary phenomena*, 1–34. Berlin: Language Science Press. DOI: 10.5281/zenodo.777526.
- Selkirk, Elisabeth. 1984. *Prosody and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Shattuck-Hufnagel, Stefanie & Alice E. Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2). 193–247. DOI: 10.1007/bf01708572.
- Singmann, Henrik, Ben Bolker, Jake Westfall & Frederik Aust. 2015. *afex: Analysis of factorial experiments*. R package version 0.13–145. <https://CRAN.R-project.org/package=afex>.
- Stevens, Kenneth N. & Samuel Jay Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65(1). 81–106. DOI: 10.2307/414843.
- Tabain, Marija. 2003. Effects of prosodic boundary on /aC/ sequences: Articulatory results. *The Journal of the Acoustical Society of America* 113(5). 2834–2849. DOI: 10.1121/1.1564013.
- Torreira, Francisco & Mirjam Ernestus. 2011. Realization of voiceless stops and vowels in conversational French and Spanish. *Laboratory Phonology* 2(2). 331–353. DOI: 10.1515/labphon.2011.012.
- Turk, Alice E., Satsuki Nakai & Mariko Sugahara. 2006. Acoustic segment durations in prosodic research: A practical guide. In Stefan Sudhoff, Denisa Lenerová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schließer (eds.), *Methods in empirical prosody research*, 1–28. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110914641.1.

- Turk, Alice E. & James R. Sawusch. 1997. The domain of accentual lengthening in American English. *Journal of Phonetics* 25(1). 25–41. DOI: 10.1006/jpho.1996.0032.
- Turk, Alice E. & Stefanie Shattuck-Hufnagel. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics* 28(4). 397–440. DOI: 10.1006/jpho.2000.0123.
- Turk, Alice E. & Stefanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics* 35(4). 445–472. DOI: 10.1016/j.wocn.2006.12.001.
- Vigário, Marina. 2003. *The prosodic word in European Portuguese*. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110900927.
- Vogel, Irene, Angeliki Athanasopoulou & Natália Brambatti Guzzo. 2018. Timing properties of (Brazilian) Portuguese and (European) Spanish. In Lori Repetti & Francisco Ordóñez (eds.), *Romance languages and linguistic theory*, 325–340. Amsterdam: Benjamins. DOI: 10.1075/rllt.14.19vog.
- White, Laurence. 2002. *English speech timing: A domain and locus approach*. University of Edinburgh. (Doctoral dissertation). <https://www.cstr.ed.ac.uk/projects/eustace/fulltext/front.pdf>.
- White, Laurence, Silvia Benavides-Varela & Katalin Mády. 2020. Are initial-consonant lengthening and final-vowel lengthening both universal word segmentation cues? *Journal of Phonetics* 81. 10098. DOI: 10.1016/j.wocn.2020.100982.
- Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America* 91(3). 1707–1717. DOI: 10.1121/1.402450.
- Winitz, Harris, M. E. Scheib & James A. Reeds. 1972. Identification of stops and vowels for the burst portion of /p, t, k/ isolated from conversational speech. *The Journal of the Acoustical Society of America* 51(4B). 1309–1317. DOI: 10.1121/1.1912976.
- Xu, Yi. 2010. In defense of lab speech. *Journal of Phonetics* 38(3). 329–336. DOI: 10.1016/j.wocn.2010.04.003.