Stony Brook University

The phonetics of stress manifestation:
Segmental variation, syllable constituency and rhythm

A Dissertation Presented

by

MI RAN KIM

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Linguistics

Stony Brook University

August 2011

Stony Brook University

The Graduate School

MI RAN KIM

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Marie K. Huffman – Dissertation Advisor
Associate Professor, Linguistics Department

Lori Repetti - Chairperson of Defense
Associate Professor, Linguistics Department

José Elías-Ulloa
Assistant Professor, Hispanic Language & Literature

Lisa Lavoie
Emerson College

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

ii

Abstract of the Dissertation

The phonetics of stress manifestation:

Segmental variation, syllable constituency and rhythm

by

MI RAN KIM

Doctor of Philosophy

in

Linguistics

Stony Brook University

2011

This dissertation investigates how lexical stress in Spanish is phonetically manifested within the syllable, and how stress manifestation interacts with other prosodic effects, such as phrasal accent and initial word-boundary effects. One of the challenges in the study of prosody is the fact that multiple prosodic factors induce variation in a limited number of phonetic dimensions. For this reason, results in the literature have often been inconclusive or conflicting. The disentanglement of these prosodic effects will help us understand the complex nature of prosody.

The results of this study show that stress effects can be identified independently of phrasal level prominence (accent) and prosodic-boundary effects (position in word). Among the prosodic effects, durational and spectral properties clearly differentiate stressed onset and vowel from their unstressed counterparts, but intensity does not. In addition, similarities between stress and accent effects, grouped as 'prominence effects', are distinguished from word-level boundary effects that are conditioned by accent conditions. Finally, the results indicate that syllable constituency can be involved in prosodic manifestation, possibly constrained by higher-level prosody, such as isochrony (stress-timed vs. syllable-timed rhythms). We identified apparent complementarity between syllable constituents in the stress manifestation of two dialects of Spanish, which can be

attributed to two structural motivations. On the one hand, syllable constituency is involved in segmental variation at the lexical-level (stress); and on the other hand, a higher-level prosody (isochronous rhythm) can play a role as an upper bound in constraining the variability. A gestural account is provided to accommodate the asymmetric contribution of syllable constituents to the temporal manifestation of stress.

The results on prosodic effects reported in this dissertation provide several linguistic implications that are important for future research. Firstly, syllable prominence may be expressed by relative phonetic values (such as C/V ratio in duration and intensity) between syllable constituents. Secondly, the word boundary effect is different from prominence effects (stress and accent), and it may even be conditional to the prominence factors. Cross-linguistic data in the future will allow us to determine whether conditioning nature of the boundary effect may be language universal. Thirdly, temporal modulations within a prosodic domain (i.e., syllable) can vary even at the level of dialects. The narrow measurement method provided insight into how sub-segmental units can be affected by prosody, and can be linguistically encoded in the manifestation of prosody. The final implication is related to the question of why prosodic effects are rarely phonologized. Prosodic effects are gradient but not always cumulative; the effects partly depend on which prosodic factors are involved and how the multiple prosodic factors interact in a specific phonetic dimension. Thus, the interrelations among prosodic factors are one of the key questions in answering what the role of prosody in the grammar can be.

The investigation of the phonetics of prosody manifestation provides us insight into interrelations among prosodic factors, and the systematicity behind the complexity of prosodic effects. It is hoped that this dissertation contributes to a better understanding of prosody as a well-structured grammatical component, and its connections to segmental phenomena in languages.

To my family

나의 사랑하는 부모님,

든든한 두 언니와 남동생에게

이 논문을 바칩니다.

TABLE OF CONTENTS

List of Figures

x

List of Tables

List of Abbreviations

Suprasegmental	IP	Intonational Phrase
Descriptions	PoW	Position in Word
	PoIP	Position in IP
	WI	Word-initial position
	WM	Word-medial position
	Str.	Stressed condition
	Unstr.	Unstressed condition
	Acc.	Accented
	Unacc.	Unaccented
Prosodic index	P	Position in word ([+ PoW] = word initial)
	S	Stress ([+STRESS] = stressed)
	A	Accent ([+ACCENT] = accented)
Language/dialect	Sp.	Spanish
	AS	Argentinian Spanish
	PS	Peruvian Spanish
	AE	American English
	BE	British English
Segmental	C	Consonant
Description	V	Vowel
	Syll.	Syllable
	T_n	Transition (T_1~T_4)
	Css	Steady-state of consonants
	Vss	Steady-state of vowels
Acoustic	COG	Center of Gravity
Measurement	M1	First spectral moment (Center of Gravity)
	M2	Second spectral moment (Variance)
	M3	Third spectral moment (Skewness)
	M4	Fourth spectral moment (Kurtosis)
	RMS	Root Mean Square
Applications	TADA	Task Dynamics Applications

Acknowledgments

And now, I have to agree with the world saying that being stubborn is not a great personality trait to have, especially in academia. Admitting this, I feel myself so lucky to have the teachers and friends who helped me to transform "being stubborn" into the strength by which I was possibly able to complete this dissertation.

First and foremost, I would like to thank my dissertation advisor, Prof. Marie Huffman, who has not only guided me with every single step necessary to complete this dissertation (except writing this page), but also showed me why and how to see "open doors". Without her, I might have still been standing there looking at the closed ones. I want to thank my dissertation committee members, Prof. Lori Repetti, Prof. José Elías-Ulloa, and Dr. Lisa Lavoie for their helpful discussions and encouragement. I feel again very lucky to work with the committee who gave me heartfelt friendship as well as academic inspiration. I would also like to thank my teachers at Korea University: Prof. Keeho Kim, who introduced me to a world of linguistic melody, and Prof. Jae-Woong Choe, who encouraged this humble student to become a linguist. I would also like to warmly acknowledge the academic benefits and the support that I have received from the Department of Linguistics at Stony Brook. My gratefulness for all my teachers will stay with every moment of success in my academic career.

I am very grateful also for the warm support and friendship that I have experienced during my years at Stony Brook. I want to thank Dr. Joy Janzen, who has inspired me to appreciate the beauty of the world outside linguistics (but still in the world of teaching). I want to thank also my beloved and invaluable colleagues and friends: Susana Huidobro, who helped me to do "TA work" successfully in my first semester and now to live my (mental and physical) life healthy; Yu-an Lu, an Angel of mine, who is always there to help and support me; Emily Romanello, the shared joy of friendship between us will last forever in my memories of Stony Brook campus. We often say things without doing them, but you have been doing good things for others without need of saying. Thank you for being you. I owe also so many things to my

current and former colleagues at Stony Brook for the invaluable academic/social interactions and the unforgettable moments that we have shared. Too many to name here: so, let me thank you via Facebook.

My special thanks go to Dr. Hosung Nam, who was a great colleague, a good friend, an insightful (informal) advisor for my master's thesis and also this dissertation, and a wonderful collaborator. I owe you so much and I will pay you back with good research in the future. I also want express my thanks to my informants who were participated in the data recording. I wish you all great academic achievements as they complete their studies at Stony Brook.

Last, I wish to thank my family for their endless love and support, without which no accomplishments can be appreciated. My mother, father, sisters and brother: I have been so far away and too long from home, but you have always been with me. I also thank Hijo Kang, one of my beloved colleagues, and his family for their generosity and friendship, in which I could feel home here. Soonhui Choi, and the two little big boys, Wongyu Kang, and Hyeonggyu Kang, I thank you for your smiles in love. Your patience, understanding and encouragement made all the difference.

Chapter 1 Introduction

This dissertation examines how lexical stress is phonetically manifested within the syllable and how stress manifestation interacts with other prosodic components including phrasal level prominence (accent), position in word (PoW), and rhythmic regularity, all of which overlap in the phonetic realization of segments. The goal of this study is to identify the prosodic effect that is specific to stress, and to disentangle the stress effect from other prosodic effects. Specifically, the dissertation addresses three main research questions: [1] How does stress affect the phonetic manifestation of the syllable in relation to the syllable constituency (onset and vowel)? [2] How does the stress effect interact with other prosodic effects such as accent and position in word effects? [3] How is rhythmic regularity in a language involved in such prosodic interactions? These research questions are addressed through phonetic investigations of prosodic manifestation patterns in two dialects of Spanish (Peruvian and Argentinian),

The first research question involves phonetic investigation of stress effects. Lexical stress is one of the conditioning factors that regularly influences segmental variation. For example, English full vowels are preserved in stressed syllables, but may undergo some degree of centralization/reduction when unstressed, often resulting in schwa (Chomsky & Halle 1968). Similarly, in Brazilian Portuguese, stress is involved in phonemic vowel distribution, allowing a limited vowel set in unstressed syllables (Major 1985): among the 8 phonemic monophthong vowels, [ɛ, ɔ] become neutralized to [e, o] pretonically, which are further raised to [i, u] posttonically. Restricted distribution phenomena conditioned by stress are common across stress-system languages, and they are not limited to vowels. Allophonic distribution of consonants also refers to prosody. For instance, the distribution of English aspirated voiceless stops in onset depends on stress and/or position in word: aspirated voiceless stops [pʰ, tʰ, kʰ] in English can occur in the onset of stressed syllables or in word-initial position. An experimental study by Lavoie (2001) shows that consonant realizations are greatly influenced by stress (and position in word) in both duration and even manner of articulation.

Variable consonant realization under stress has also been reported in the literature, mostly within a language (Gordon 1997; Sproat & Fujimura, 1993; Dalcher, 2006; among others), and stress related variation seems to be language-specific. In English, for instance, aspirated stops occur in stressed (or word initial) onsets and flapping in unstressed positions; on the other hand, German does not have flapping in the same prosodic context. Lavoie (2001) investigated inter-language variation with consonants in English and Spanish. Her results showed that the realization of consonants in both languages has consistently varying patterns sensitive to stress (and position in word): specifically, both duration and magnitude in articulation of consonants are sensitive to stress, the effects of which vary in degree by segment type, and also by language. Besides prosodic contexts, the phonetic realization of a given segment can be affected by non-prosodic factors such as phoneme inventory (Manuel 1990; Lavoie 2001). Speakers make efforts to preserve linguistic contrast (e.g., segmental inventory) in their production and realization of segments can be constrained by this inventory effect. Such variation needs to be further studied to understand how it relates to linguistic knowledge.

Although it is clear in the literature that both onsets and vowels are affected by stress, it is not clear whether the two interact. For example, a syntagmatic relation between the vowel and the coda within the syllable is relatively well-known (e.g., compensatory lengthening between the constituents: Hayes 1989; Broselow et al. 1997), but a parallel relation is rare phonologically between onset and vowel. Our question is whether there is such a similar relation in the phonetic dimension between the onset and the vowel, both of which constitute the syllable and possibly contribute to making syllables phonetically prominent. Assuming that the syllable is the smallest domain for stress manifestation (e.g., stressed and unstressed syllables as being minimally contrastive in terms of prominence), it seems likely that there could be complimentarity in stress effects between subparts of the syllable.

The second question addresses possible interactions and/or interrelations among prosodic factors. Recognition of the importance of prosody has led to a number of instrumental studies aimed at discovering phonetic correlates of prosody. Since stress is realized in multiple phonetic dimensions, including time, intensity, and pitch, results reported in the literature have not always agreed on which cues correlate with lexical stress. One of the reasons underlying these

mismatching results is that the term of "stress" was used to represent any kind of prominence carried on the syllable lexically. For example, duration, intensity, and pitch has often been reported as the phonetic/perceptual correlates of English stress (Fry 1955, 1958), although the independent contribution of each parameter to perception has been variably reported in the literature (Fry 1955, 1958; Lieberman 1960; Sluijter & van Heuven 1996a; Turk & Sawusch 1996). As many researchers have recently pointed out, stress correlates are very often the combinatory result of stress with other levels of prosody such as pitch accent, intonation, or focus (Beckman 1986; Sluijter & van Heuven 1996b; Cho & Keating 2007). This is mainly because of another intrinsic characteristic of prosody; that is, the phonetic realization of multiple levels of prosody shares physical dimensions. Thus, discrepancies found in the literature on stress correlates are due at least in part to the fact that stress has been confounded with other prosodic factors such as pitch accent or prosodic boundary adjacency. This may be why, as Lehiste (1970: 106) says, '… stress has for a long time been the most elusive one [among suprasegmental features]." In this dissertation, the phonetic manifestations of the three prosodic factors (stress, accent, and position in word) are examined to disentangle individual prosodic effects and/or possible interactions among them.

The third research question is primarily concerned with how rhythmic characteristics of a language are involved in prosodic interactions. The gestural undershoot hypothesis (Lindblom 1963), says that temporal adjustment by prosody may in turn induce variable segmental realizations (e.g., consonant lenition or reduced vowel). What this suggests is that lower level constituents are restricted or conditioned by the prosody at higher levels. Whether the higher level is a syllable, word, or phrasal unit, corresponding lower level constituents can be affected. In particular, isochronous prosodic patterns in a given language may play a similar role in constraining temporal aspects of segmental variation. Following the view of isochronicity as a gradient feature (e.g., Delattre 1964; Crystal 1969; Dauer 1983; Bertinetto 1989), this dissertation examines how isochronicity plays a role in prosodic interactions.

The three research questions taken up in this dissertation will be addressed through phonetic examination of prosodic effects in temporal, intensity, and spectral dimensions. We examine specifically the durational, amplitudinal, and spectral changes produced by prosodic

factors in onsets, vowels, and/or entire syllables, respectively. Based on these phonetic results, we aim to separate stress effects from other effects in each phonetic dimension (temporal, intensity, and spectral), and to characterize phonetic modifications in segments specific to either individual or combinatory prosodic effects (stress, accent, position in word). In addition, we are interested in how the onset and the vowel contribute to making syllables phonetically prominent in prosodically strong positions (stressed, accented, word-initial syllables). Our phonetic investigations will make use of acoustically available sub-phonemic information (e.g., steady-state and transitions) as well as traditional notions of segmental boundaries (e.g., C and V), which represent abstract phoneme units. Methodologically, we will take a cross-dialectal approach in investigating prosodic effects examining two dialects of Spanish. A cross-dialectal comparison enables us to maximize the information gained about prosodic effects by minimizing phonological differences that are known to affect segmental variation.

The next four subsections establish related background; they are structured as follows. Section 1.1 reviews some of the important literature on effects of prosodic factors, including lexical stress, phrasal accent, word level boundary, and sentence level rhythm as focus on the isochronous characteristics of the selected language (Spanish). In section 1.2, we discuss non-prosodic factors such as phoneme inventory and syllable constituency that are known to contribute to segmental variability in languages. Section 1.3 introduces the language and dialects investigated in this study. Section 1.4 elaborates the research questions to be empirically tested in the current study. The structure of the dissertation is summarized in section 1. 5.

1.1 Prosodic components involved in segmental variation

There are phonological components that potentially influence segmental variation. In this section, we will consider three prosodic factors that are involved in segmental variability, and also examine rhythmic regularity as one of the higher-level prosodic components that can influence stress manifestation.

1.1.1 Lexical stress effects

Commonly known stress effects are variation in timing (duration), amplitude (loudness) and F0 (pitch). Stressed vowels tend to be longer in duration, greater in amplitude and higher in pitch (Lehiste, 1970). Beyond this general description of stress effects, many studies have shown that different languages linguistically encode different phonetic correlates of stress. English vowel durations are influenced by stress and listeners attend more to duration as a main cue in the perception of stress (Klatt 1976; Crystal & House 1988). Independent contributions of each parameter to perception have been variably reported in the literature (Fry 1955, 1958; Lieberman 1960; Sluijter & van Heuven 1996a; Turk & Sawusch 1996). Lieberman (1960) found that pitch is the most reliable stress correlate followed by intensity and then duration. On the other hand, Fry (1958) places a durational cue higher than intensity revising his earlier result where pitch was the most important correlate of stress. He describes pitch effects on perception as "all-or-none," meaning that stress identification is altered by the presence of fundamental frequency change itself, and not by the magnitude of fundamental frequency change. When a fundamental frequency change influences the perception of the listeners, it outweighs the effects from both duration and intensity. Duration as a most reliable cue to stress is also found in Turk & Sawusch (1996) and Sluijter & Heuven (1996b).

A number of instrumental studies in the literature have discovered multiple and sometimes contradicting phonetic correlates of prosody. One of the reasons underlying these mismatching results is that the notion of "stress" was used to represent any kind of prominence carried on the syllable identified as being lexically stressed. What is defined as stress effects is very often the combinatory result of stress with other levels of prosody such as pitch accent, intonation, or focus (Beckman 1986; Sluijter & van Heuven 1996b). For example, Lieberman (1960) found that pitch is the most salient cue to stress for American English listeners. He used noun-verb pairs embedded in a sentence, where a noun was placed sentence finally while verbs were medial: 'We had a *cóntract*' vs. 'Don't *contráct* the flu'. Since a sentence-finally located noun in English is a potential landing site of nuclear pitch accent, the target words in the experiment could have been confounded with a nuclear pitch accent more for nouns than for

verbs. Considering the fact that the nuclear pitch accent is also most prominent in the given sentences, and pitch range expansion is found to be the greatest in the declarative statement tune (L+)H* L-L% in English (Ayers 1996), it is predictable to some extent why Lieberman's (1960) results showed the F0 cue as most influential in his experiments as opposed to the results from other studies. That is, most of the stressed syllables (e.g, '*con-*' in '*contract*' as a noun) are produced in a nuclear pitch accent position, and the unstressed syllables come from verbs that are located in sentence medial position, which is often associated with a pre-nuclear pitch accent. This means that the pitch prominence as one apparent correlate of stress cannot be solely attributable to stress effects; instead, the pitch prominence can be either a combinatory effect between stress and accent or just an accent effect. Thus, closer examination of prosodic effects is needed to evaluate the conclusions about stress correlates discussed in previous literature.

This point was also emphasized in Cambier-Langeveld & Turk (1999). They examined accentual lengthening in Dutch and English, controlling position in word, and found that the results were indeed similar, unlike what had been previously reported. Given the fact that prosodic effects can be overlapping in the phonetic realization of segments, and that they can also have different manifestations, the question is whether the individual contributions of different prosodic factors can be identified. This question was explicitly addressed in Cho & Keating (2007) for English. They found that boundary effects were differentiated from prominence effects (stress and accent). Specifically, the initial boundary affected mainly consonants (e.g., VOT) while stress/accent effects influenced mainly vowels. Since domain-initial strengthening also seemed to affect the vowels to some extent (e.g., increasing intensity), they concluded that domain-initial strengthening is not strictly local to the initial segment, but it can rather be a gradient effect. They also examined whether the domain-initial effect interacts with prominence effects (stress and accent), and found that initial segments employed the greatest intensity when no prominence is involved in the initial segment (unstressed-unaccented-initial segment). The effect was the least when other prominence effects were concomitantly present. Their interpretation of the results was that prominence effects (stress and accent) might constrain the domain-initial effect due to a possible ceiling effect. This implies that boundary effects can depend on stress and/or accent effects.

Thus, it is important to investigate stress effects independently of other prosodic effects since effects from different levels can be confounded in time; e.g., '*con*' in *cóntract* in English carries prosodic information of stress, as well as accent and/or word-initial effects. The disentanglement of prosodic effects will help us understand the complex nature of these effects.

1.1.2 Accent effects as phrasal level prominence

The terms "stress" and "accent" are often used synonymously in the literature, sometimes resulting in inconsistent (or even incompatible) observations and conclusions. Here we use 'stress' as prominence at the lexical level while 'accent' as prominence at the phrasal level. The two types of prominence are encoded in the intonational and/or semantic/pragmatic structure of a given phrase. For example, Beckman (1986: 1) defines them distinctly as follows:

> "… Accent means a system of syntagmatic contrasts used to construct prosodic patterns which divide an utterance into a succession of shorter phrases, and to specify relationships among these patterns which organize them into larger phrasal groups. And stress means a phonologically delimitable type of accent in which the pitch shape of the accentual pattern cannot be specified in the lexicon but rather is chosen for a specific utterance from an inventory of shapes provided by the intonation system…'

Following Beckman, stress is defined for the purpose of the current study as a linguistic property that has a contrastive function at the word level. Accent is another linguistic entity whose role can be distinct from stress, although the realization of the two is often not independent; that is, most accents are associated with stressed syllables.

It has been argued in the literature that pitch is a main correlate of (phrasal) accent (Bolinger 1961; Navarro 1964; Bolinger 1972; Beckman 1986; Ortega-Llebaria 2006). For instance, Ortega-Llebaria (2006) investigated the phonetic correlates of accent and stress in Spanish, and found that both stressed and unstressed syllables displayed a flat intonation contour when unaccented (post-focal position). What this means is that accent as phrasal level prominence can override the effects of lower level prominence such as lexical stress. In addition

7

to stress, other phonetic modifications are involved in accentuation. Eefting (1991), for example, showed that all the segments (and syllables) contribute to the durational changes affected by accentuation in Dutch. In addition, stress effects can be present regardless of accent conditions in the same language (Nooteboom, 1972). Other studies further confirmed that stress and accent effects differ in the phonetic dimension: e.g., overall intensity is a correlate of accent rather than stress (Sluijter & van Heuven 1996a). Thus, stress and accent are two independent sources of prosodic effects on segments.

1.1.3 Position in word as prosodic boundary strength

The third prosodic factor we are concerned with is position in word, which can be viewed as a word-level boundary effect. While the first two factors can be defined as metrically structured prominence, position in word effects are generally viewed as boundary adjacent local prominence.

The initial segment/syllable in a given word is considered a prosodically strong position, both phonetically and phonologically, as well as psycholinguistically (Smith 2002). Phonetically, initial segments undergo temporal and/or spatial strengthening (Fougeron & Keating 1997; Lavoie 2001; Byrd & Saltzman 2003; Keating et al. 2003). Phonologically, certain phonological rules are either applied or blocked; e.g., in initial position, positional neutralization is rare (Barnes 2006) and also place assimilation is unusual in that position. The phonetic/phonological strength of initial position is in line with the functional role of the initial position in word recognition examined in the word-recognition and perception literature (Lehiste, 1965; Quené, 1989; Marslen-Wilsen & Zwitserlood 1989; Davis et al. 2002; Salverda et al. 2003). However, this apparent privilege of initial position is not a phenomenon that we can take for granted. Detailed investigations have revealed that there is a great deal of variation in the phonetic realization of segments depending on other linguistic factors; e.g., domain of the effect (C, V, or syllable), number of syllables in the word (greater effect on monosyllable word), prominence (stress or accent effects), and multiple sources of effects (duration and/or spatial strengthening). Among

the factors, we are interested in how stress affects each syllable constituent (C, V, or syllable as a whole).

While the boundary effect in general lengthens the vowel as well as the initial consonant (Klatt 1975; Ladd & Campbell 1991; Wightman et al. 1992), Fougeron & Keating (1997) found no effects on the vowel for English and concluded that the strengthening is "a localized effect at prosodic domain edges".[1] This locality hypothesis is also supported by Cho & Keating (2007): an onset C shows articulatory strengthening (temporal and spatial effects) but following vowels have no strengthening effects. More importantly, Cho & Keating (2007) investigated effects of prosodic factors (utterance-initial strengthening, stress, and accent) in English, finding that word-initial consonants are sensitive to domain-initial strengthening while stress and accent mostly affect vowels. In addition, the boundary effect can be conditioned by stress or accent. This suggests that prosodic manifestation varies within the syllable, influencing consonant and vowel differently. Thus, position in word effect also needs to be investigated in relation to stress and accent effects.

1.1.4 Rhythmic regularity and its role in segmental variability

The fourth prosodic component that is hypothesized to be relevant in this dissertation is rhythmic regularity, often referred to as "isochrony" (Trubetskoy 1938; Pike 1945; Abercrombie 1967; Dauer 1983). Pike (1945) suggested that rhythms of languages can typologically classified referring to the dichotomy between "syllable-timed" and "stressed-timed" languages.[2] The stress-timed rhythm refers to timing pattern with equal duration between stressed syllables, while the syllable-time rhythm with equal duration between syllables.

---

[1] Fougeron (2001) also shows that the local effect is also the case in French.

[2] The earlier idea of isochronous linguistic rhythm is also found in Trubetskoy (1938). The term "isochrony" is used based on the general idea of how rhythmic characteristics among languages vary. It is not to assume that stressed and unstressed syllables are isochronous. Instead, it is to assume that the durational difference between stressed and unstressed syllables is relatively smaller than others.

In the literature, some studies have attempted to provide ways either to determine this rhythmic dichotomy (syllable-timed vs. stress-timed) phonologically or to measure it. For example, a number of phonological criteria have been suggested based on typological behaviors of languages that may represent the two opposing types of isochrony (Bertinetto 1989). These include vowel quality (vowel reduction vs. full articulation) of unstressed syllables; consonant quality (more variable vs. similar articulation) in unstressed syllables; native speakers' knowledge of syllable boundaries (relative certainty vs. uncertainty) in counting syllables; temporal acceleration (e.g., speaking rate) mainly obtained from compression of unstressed syllables vs. proportional compression.[3] These properties seem to describe the characteristic differences between, for example, English/German and Italian/Spanish (Bertinetto, 1989). However, experimental studies have failed to confirm such criteria in grouping languages into the two categories. Roach (1982), for instance, tested the hypothesis of Abercrombie (1967) with the assumption that syllable length is greatly varied in stress-timed and equal in syllable-timed languages. The overall results of Roach showed that the syllable variability measured by standard deviation of syllable durations tends to be greater in the stress-timed languages than in the syllable-timed, but the results could not categorize all the languages investigated. In addition, there appeared either overlapping or contradictory results among the languages.

Although empirical studies failed to support isochrony as a categorical division, this does not disprove the existence of isochronous rhythmic characteristics among languages, as found in a number of linguistic behaviors. For instance, infants can distinguish the two types of rhythm (Nazzi et al. 1998; Nazzi & Ramus 2003), and adult learners make use of regularities found in syllables or feet in processing speech (Cutler et al. 1992). In addition, second language learners show their native knowledge of rhythm in switching from native to second language production (White & Matthys 2007). As Dauer (1983) suggests, the two isochrony categories may not be mutually exclusive, and they are rather gradient in nature. This is interpreted to mean that a

---

[3] Other patterns include: complex syllable structure with relatively uncertain syllable boundaries vs. simple syllable structure with well-defined syllable boundaries, position of stress (flexible vs. fixed), and so forth. More criteria have also been proposed in Schmid (2004), including the preference for closed syllables in stressed-timed languages.

language may bear more or fewer stress-timed characteristics than syllable-timed ones, but not be necessarily affiliated with one of the two categories. The concept of "isochrony" is taken in this dissertation as rhythmic regularities that can be gradient among languages rather than as exclusive rhythmic categories.

Within the prosodic hierarchy (e.g., Selkirk 1978, 1982; Nespor & Vogel 1986), it is assumed that higher-level prosody (e.g., Utterance and Intonational Phrase) is projected from lower-level prosodic components (e.g., Mora and Syllable). In addition, Selkirk (1984, p.12) suggests that, "the metrical grid alignment of a sentence is a representation in terms of which such things as the isochrony of stressed syllables and more generally the relative durations of syllables might be expressed". This can be interpreted to mean that prominence at various levels may result from hierarchical relations (or interactions) at lower levels. This prosodic connection is also supported by the following phonetic/phonological phenomena, all of which can be language-specific. Firstly, phonetic realizations of two lexically prominent syllables (primary stresses) can differ depending on the relative strength of phrasal accents (e.g., nuclear vs. non-nuclear pitch accent) associated with the syllable (Palmer 1922; O'Connor & Arnold 1961; Crystal 1969). Secondly, at the word-level, the degree of prominence under primary stress differs from secondary stress, and thus segmental realizations are affected according to the relative strength of stress. For example, in Chickasaw, duration and intensity differ from primary and secondary stress realization (Gordon 2004), and gradient lengthening (or strengthening) differs between primary and secondary stress in Finnish (Suomi & Ylitalo 2004).[4] This again means that the manifestation of the syllable as a unit is bounded by the relative strength of the stress assigned to the syllable. Finally, rhythm characteristics including speech tempo modify segmental realizations. For example, languages show temporal adjustments possibly depending on the number of segments in a syllable (e.g., closed-syllable shortening) or number of syllables in a word (e.g., word compression effects).

---

[4] Phonetic correlates of secondary stress are mainly in vowel formant (F1) and weakly in pitch (cited in Gama Rossi 1998; Abaurre & Fernandes-Svartman 2002).

As mentioned earlier in this chapter, the gestural undershoot hypothesis (Lindblom 1963) predicts that temporal adjustment by prosody may in turn induce variable segmental realizations (e.g., consonant lenition or reduced vowel). That is, lower-level constituents can be restricted or conditioned by the prosody at the higher-level: whether the higher-level is a syllable, word, or phrasal unit, lower- level constituents will be affected. Similarly, it should then be possible that isochronous characteristics at a higher level of prosody in a given language can also play a role in constraining segmental variation. For example, Solé (1991) takes a view of isochrony as one of the prosodic constraints on segmental variation in English. Following the view of isochronicity as a gradient feature (e.g., Delattre 1964; Crystal 1969; Dauer 1983; Bertinetto 1989), Solé argues that the existence of weak forms in the absence of stress can be "a consequence of the stress-timed and isochronous nature of English rhythm" (Solé 1991: 155). Further investigations are needed for a better understanding of the involvement of isochrony as a prosodic constraint on segmental variation. Our investigations in the current study include this isochrony hypothesis and test whether isochronous rhythm in Spanish plays a role in manifesting stress within the syllable. In particular, we are interested in whether and how the stress effect on the syllable onset in the two dialects varies with respect to the stress effect on the vowel. If there is a complementary relationship between syllable constituents, it can be interpreted to mean that the syllable constituents interact with each other to preserve the upper level prosodic constraint, isochronous rhythm.

Spanish is often described as a syllable-timed language, because it shows a *relatively* regular time interval between syllables regardless of stress (Pike 1945; Hockett 1958; O'Connor 1973; Amador-Hernandez 1986; Solé 1991). Solé (1991), for example, identifies Spanish as a syllable-timed language taking multiple criteria often used in the literature. Three of the criteria are as follows: (1) unstressed syllables show vowel reduction due to the amount of time allotted to every syllable. (2) The way speaking rate influences syllables differs between English and Spanish, such that speaking rate effects on the durations of both stressed and unstressed syllables are roughly the same in Spanish, whereas the effect is not proportional between stressed and unstressed syllables in English. (3) The syllable is the rhythmical unit in Spanish verse (e.g., 'octosyllable' or 'decasyllable' used in describing verse) while it is the stress (or metric system:

iambic, trochee, etc.) that occurs regularly in English verse. Thus, the degree to which English and Spanish differ in stress manifestation may be understood such that Spanish is "closer to" a syllable-timed pole in the extreme, while English closer to the other side, "stress-timed".

Though it may not be necessary, it is possible for Spanish to refer to its overall isochronous rhythm as a prosodic upper boundary in making stressed syllables prominent. We will consider this possibility of rhythmic effect in interpreting the phonetic results of stress effects in relation to syllable constituents (onset and rhyme vowel).

1.2 Non-prosodic components affecting segmental variability

Literature has also shown that variability of segments can be constrained by non-prosodic factors such as phoneme inventory and syllable constituency. We will discuss each of these briefly in turn.

1.2.1 Phoneme inventory effects and variability

The first non-prosodic factor we will consider is phoneme inventory (Manuel 1990; Lavoie 2001). Speakers make efforts to preserve linguistic contrast (e.g., segmental inventory) in their production, even in the lower-level of modification regulated by certain contextual effects. The realization of segments can be constrained by this inventory effect. Manuel (1990) showed that degree of coarticulation (V-V sequences) can vary depending on vowel inventory. For example, Shona, Ndebele, and Sotho differ in their low vowel distributions: Sotho has a more crowded non-high vowel space with 7 phonemic vowels (/i, e, ɛ, ɑ, ɔ, o, u/), whereas Shona and Ndebele have five vowels (/i, e, ɑ, o, u/). It was found that Sotho /ɑ/ shows les coarticulation than do the other two languages, which do not have nearby neighboring vowels phonemically contrasting with /ɑ/. Manuel attributed the varying degree of the vowel variation (coarticulation) among the languages to the phonetic space that each language uses to encode contrast. The idea that segmental realization could be determined by the need to preserve phonological distinctions is

13

consistent directly and indirectly with various results in the literature (Öhman 1966; Manuel & Krakow 1984; Jongman et al. 1985; Flemming 2001; Lavoie 2002). This inventory effect is not limited to vowels. Lavoie (2002) investigated subphonemic effects on consonant variation, and she explained the variable degree of consonant realization on the surface by reference to the idea that inventory differences allow different kinds of space for allophonic variation. Specifically, comparing phonetic properties of consonant variation in English and Spanish, she found that Spanish /k/ varies less than English /k/. She attributed the variability difference to the existence/absence of /x/ in Spanish/English inventory, respectively. Table 1.1 (b) illustrates the results of Lavoie (2002) showing the differences in /k/ variation in the two languages.

Table 1.1 Sub-allophonic manner variation result of /k/ in English and Spanish (Lavoie 2002)

(a) Velar obstruent inventory

	English	Spanish
Stop	k	k
Fricative	-	x

(b) Non-stop realization of /k/

	English	Spanish
Fricativization	13%	3%
Approximant	8%	4%

In Table 1.1(a), velar obstruents in the two languages are summarized. The voiceless velar fricative is present in Spanish but absent in English. Lavoie's results show that greater variation is found in English for the voiceless stop /k/ compared to Spanish. Lavoie's (2002) interpretation is that the inventory difference for velar obstruents plays a role in the realization of /k/, and this is why velar stops in English appear to vary more than Spanish velar stops. As a result, more variable sub-allophonic realization of this stop is predicted in English than in Spanish.

To isolate inventory effects from stress as an independent factor affecting segmental variation, we would ideally control inventory effects. There are two possible such cases: one is to compare two cognate languages that preserve identical phoneme inventories, and the other is to look at dialectal variation within one language. The former situation is rare or absent since even related languages have undergone numerous sound changes that are reflected in their phonological systems. For example, contemporary English and German have much in common phonologically but the two languages' consonant inventories are not strictly comparable. In this case, a global investigation of stress effect may not be guaranteed since the inventory effect can

obscure stress effects. If we consider both consonant and vowel in the same study, the chance of finding matching inventories between languages becomes even lower. Therefore, the second case, a dialectal comparison within a language, would give us a fair chance to study the fine-grained phonetic details of stress manifestation patterns in terms of degree of variation and possible interactions among the phonetic parameters without the confound of inventory differences.

1.2.2 Syntagmatic relations within the syllable

The second non-prosodic factor is the domain of stress manifestation, which is the syllable, and the syntagmatic relation between its constituents (e.g., Cho & Keating 2007). We have seen so far that consonants and vowels are realized variably under the influence of sound inventories. However, since most studies on segmental variation are in fact looking at different constituents in a syllable, either onset consonant or vowel (rhyme), we only know the existence of paradigmatic effects of stress. Much is known regarding phonetic/phonological behavior of the rhyme (including coda consonants) based on syllable weight. For instance, codas can contribute to syllable weight in a weight-sensitive language system. Moraic theory accounts for the weight involvement of coda consonants by assigning a mora link to coda (Hyman 1985; Hayes 1989). In addition, this phonological structure can be reflected in the phonetic manifestation in a given language: the duration of a segment that occupies a whole mora is longer than that of a mora-sharing segment (Broselow et al. 1997). The syllable structure assumed here is shown in (1)-a (e.g., Halle & Vergnaud 1978; Selkirk 1982), and the moraic structure of light and heavy syllables are in (1)-b and c, respectively:[5]

---

[5] This binary branching structure within the rhyme is proposed also in earlier literature of syllables (Pike & Pike 1947; Kurylowicz 1948; Fudge 1969)

(1) a. Syllable structure   b. Non-moraic coda   c. Moraic coda
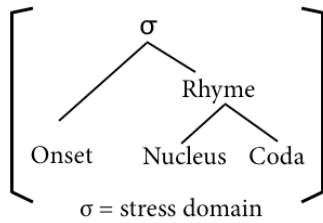
σ = stress domain
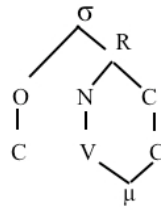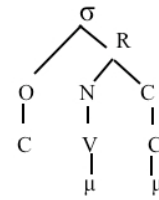
Moraic theory predicts that the syllable represented in (1)-b is light while the one in (1)-c is heavy, since the mora as a unit represents syllable weight. Constituency within the rhyme, as illustrated in (1), seems to play a role for compensatory lengthening, or iambic lengthening (Hubbard 1995; Broselow et al. 1997). For instance, Broselow et al. (1997) showed that the duration of the vowel in (1)-b is phonetically shorter than the one in (1)-c, as the phonological structures are represented in (1). The phonetic evidence of durational patterns led them to the conclusion that the phonetic timing (durational pattern) of a language reflects such underlying prosodic structures (e.g., moraic status of coda). Thus, as moraic theory implies, the behavior of coda consonant is bound to the rhyme as a whole rather than the coda itself.

On the other hand, the exclusion of the onset duration from such a calculation is probably due to the general assumption that the onset is phonologically weightless in stress-system languages. In stressed syllables, the onset is known to show a great deal of variation, as examined in Lavoie (2001). She found that onset consonants in both English and Spanish vary a lot in duration and articulatory magnitude, and that onsets even change manner of articulation; e.g., stops can be realized without closure or burst, and can become (phonetically) fricative or approximants. This variable consonant realization is greatly influenced by stress and position in word. In addition, we discussed earlier that the allophonic distribution of a consonant may refer to stress/word-initial position (e.g., English stop consonants: aspirated stops in stressed or word-initial position and the other allophones elsewhere). Thus, stress affects onset consonants as well as constituents of the rhyme.

Gesturally, it is assumed that the onset consonant and the vowel are in-phased. This means that C and V are coordinated together in time, and initiated together in gestural planning. The prediction is treated as fundamental within Articulatory Phonology (Browman & Goldstein 1989). The following diagram schematically describes this gestural coordination between C and V within the syllable (Browman & Goldstein 1989, 2000).[6]

(2) Phasing relation within syllable
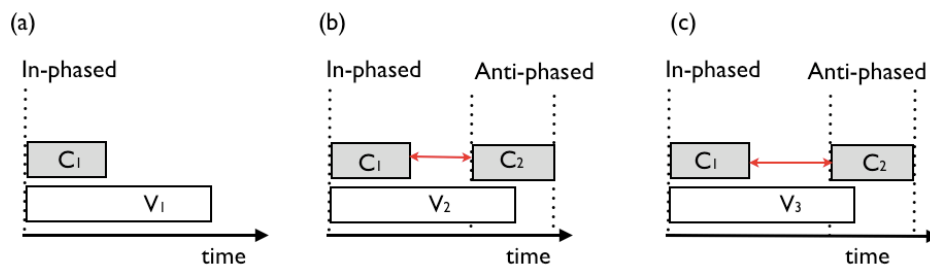
This diagram illustrates how syllable constituents can be characterized with respect to phasing relations. Given the syllable structure assumed in (1) (or the moraic structure within the syllable), variability of onset seems to be limited. For example, the onset consonant is in-phased with the vowels (a~c), and coda consonant is anti-phased, or sequential in time (b~c). In this configuration, the onset C is less likely to influence the duration of V than coda C may, given the specified time and magnitude. That is, the onset C cannot change the vowel duration unless the C gesture modifies its timing specification (e.g., the width of C_1 box). On the other hand, the coda consonant in (2b, c) is phased sequentially over V, and vowel duration of the output can change depending on where the coda C is phased: compare (b) and (c) to see that the vowel durations differ as a consequence of the C_2 phasing within the syllable. The arrow in (b) and (c) indicates the potential vowel duration on the surface. This gestural phasing relationship predicts a

---

[6] As posited in Articulatory phonology (Browman & Goldstein, 1988; among others), speech production is viewed as temporally organized and not as a sequence of discrete sounds. Particularly, onset consonants have been known to retain a relatively stable articulatory relationship with the following vowel, serving the consonant center (the C-center) as an anchoring place of vowel (Browman & Goldstein 1988; Honorof & Browman 1995; Byrd 1995). These facts demonstrate that onset and rhyme are locally timed within syllable.

potential compensatory relationship between vowel and coda consonant, as discussed in Nam (2004).

However, no prediction can be made from this configuration regarding onset variability under stress. It has been claimed that phonologically lower sonority onsets are preferred (Vennemann 1988; Clements 1990). The 'Core syllabification Principle' proposed by Clements (1990) also reflects the preference of having onset; e.g., consonants are in principle affiliated as onset first rather than as coda(s). In addition, though typologically rare, there are languages that show onset sensitivity to stress: syllables with onset are heavier than syllables without onsets (e.g., Arrernte and Banawá: Davis 1988; Ladefoged et al. 1997; Goedemans 1998), syllables with complex onsets (CC) can be heavier than those with simple onsets (e.g., Bislama: Camden 1977), and syllables with voiceless consonants are heavier than those with more sonorous consonants (e.g., Pirahã: Everett & Everett 1984).[7] The fact that stress can be sensitive to onset composition implies that quality of onset as well as its presence in syllables can have phonological importance.

We have seen a syntagmatic relation between vowel and coda consonant, which constitute the rhyme. The phonetic details of each constituent, nucleus and coda, are complementary to maintain the phonological weight as represented by the mora. However, if the syllable is the smallest prosodic domain for stress manifestation (e.g., Hayes, 1995) and serves as the domain of phonetic realization of stress, a similar question arises as to how each constituent (onset and rhyme) reflects phonetic modifications driven by stress, and whether there is any interaction between the two.

Consonant variation influenced by prosodic factors has largely been discussed in terms of prosodic domain effects, such as consonant strengthening/fortition, and weakening/lenition. Consonant variation is sensitive to prosodic position, and what we call "strong position" tends to preserve full consonant quality compared to "weak position", as shown diachronically (Hock 1991) and synchronically (Lavoie 2001). The assumptions that the syllable is the smallest prosodic domain for stress manifestation and that it is hierarchically organized as in (1), suggests a possible

---

[7] There are more languages reported to have onset sensitivity (Goedemans 1998; Gordon 2005).

relation between onset and rhyme (e.g., complementarity) within the syllable, as is the case between vowel and coda within rhyme. We hypothesize that onset and rhyme can contribute to the phonetic manifestation of stress in a complementary fashion, and more likely to do so if conditioned by higher prosody (e.g., isochronous rhythm in a given language). The current research tests this hypothesis by examining how stress affects both onset and vowel, in a syntagmatic perspective.

Possible complementarity between C and V within a syllable in relation to higher-level prosody is implied in Delattre's description of temporal characteristics influenced by stress among four languages (Spanish, English, French, and German): he notes that, "In languages like Spanish (*compared to English and German*), the role of vowel duration is expected to be reduced and that of consonant duration to be increased (*in the phonetic manifestation of stress or prosody in general*) [1966: p.34]".[8] This is interpreted to mean that stress effects on consonants may be greater in Spanish since the lengthening of vowel is less affected by stress among the languages investigated in his study. One of the distinct rhythmic characteristics Delattre might have observed is a constraint on higher level of prosody (so-called "isochrony"), which will be briefly considered in the following section.

1.3 Selected language and dialects

From the literature on prosody and segmental variation, we know that there are independent phonological factors that have to be controlled in investigating prosodic effects: phoneme inventory, syllable constituents, and levels of prominence, and isochronous characteristics. A cross-dialectal study was chosen for two specific reasons.

First, dialectal comparison enables us to look at subtle phonetic differences while keeping phonological discrepancies minimized, such as differences in phoneme inventory and prosodic system. As noted earlier, phonological discrepancies between languages may result in difficulties in disentangling stress effects from segmental differences. Second, a dialectal comparison allows

---

[8] My interpretation of the context is italicized.

us to control the influence of the prosodic system other than stress. Languages have their own characteristic prosodic system, and prosodic components between languages may not be comparable to one another. In addition, languages show sensitivity to this higher-level prosodic structure and they can differ in how the effect is manifested. For instance, Keating et al. (2003) investigated domain-initial articulatory strengthening across languages, and all languages showed phrasal or prosodic conditioning of articulation *to some degree* and *at some level*.[9] This implies that even though a domain-initial effect can be universal across languages, the effect may vary in terms of magnitude and prosodic levels to which a language is more (or less) sensitive. By comparing within language variation, we are more likely to find the segmental changes mainly affected by stress, and how the stress effect interacts with accent and/or position in word.

Even when we have two languages that are comparable in terms of overall prosodic system, phonetic correlates of stress or pitch accent can vary depending on how languages encode such fine-grained phonetic details into their system. For example, Sluijter & van Heuven (1996a) investigated phonetic correlates of stress and pitch accent in Dutch and American English (AE), and found that both languages use F0 and overall intensity to differentiate only between accented and unaccented syllables but not between stressed and unstressed syllables. They concluded that the two languages do not differ overall in distinguishing from stress to accent, but do differ in that vowel quality seems to be more important in American English.

Another study by Gordon (2004) illustrates how important it is to examine prosodic interaction at different levels in a given language. Gordon compared effects of primary and secondary stress in Chickasaw. He found asymmetric stress manifestation depending on phrasal accent. The manifestation of phrasal prominence in this language constrains the manifestation of stress such that secondary stress affects duration and intensity, but not F0, which is conditioned by the status of pitch accent. He argues that the word-level stress is outranked by phrasal-level prominence (accent), and thus the higher-level prosody (accent) constrains F0 variation in manifesting stress. In other words, F0 manifestation is allowed only for primary stress and not for

---

[9] They also found that the domain-initial strengthening varies among the speakers and segments within a language.

secondary stress. Thus, a dialectal comparison can provide a context where similar inventories as well as prosodic structure can maximally be controlled.

Two regionally distinct forms of Spanish spoken in Latin American were chosen: Peruvian and Argentinian. Interestingly, Argentinian Spanish is known to show rhythmically different prosody from other Spanish varieties spoken in Latin America. The source of variation is not yet clear but it has been argued that Argentinian Spanish is greatly influenced by language contact with Italian (Colantoni & Gurlekian 2004). Impressionistically, Argentinian Spanish is described as being more "rhythmical" in intonation, which may be due to differences in temporal characteristics such as how alternating stressed and unstressed syllables (or accented and unaccented words) are produced. Thus, a dialectal comparison would provide information about syllable-level stress manifestation.

Stress cues in Spanish have been described variably: pitch, duration and intensity (Navarro 1964 [cited in Ortega-Llebaria 2006]; Contreras 1977), or pitch and duration (Ortega-Llebaria 2006), or mainly pitch rather than duration (Bolinger 1961). From the differences of phonetic/perceptual correlates between English and Spanish, it seems reasonable to view patterns within allophonic variation under stress as language specific rather than automatic. In other words, it would not be surprising to find variable behaviors of vowels, and even consonants, under stress, from one language to another. An important question is whether the effect is attributable independently to stress, or it is in fact the combinatory effect of stress with other prosodic factors (phrasal prominence and prosodic boundary effects). Ortega-Llebaria (2006) investigated the phonetic correlates of accent and stress in Spanish, and found that both stressed and unstressed syllables displayed a flat intonation contour when unaccented (post-focal position).

Since the main focus of the investigation is prosody-conditioned variation, other sources of segmental variation should be as minimal as possible. We examine two phonemic fricatives in Spanish, /s/ and /x/. Regarding these segments, no major variation has been reported for the two consonants in onset position. Most consonant variation occurs in coda position in Spanish dialects. As in many dialects of Spanish, /s/ in coda position aspirates; e.g., busca 'nf. search' [buxka] (Canfield 1981), debuccalization of /s/ to [h] varies depending on dialects. Although not

directly related to the segments we investigate here, two cases of ongoing variation is worth noting here. First, Argentinian Spanish shows loss of assibilation in rhotics and assibilation in palatals (Colantoni 2001, 2006): [j]uvia > [ʒ]uvia 'rain', [ɹˇ]ubia > [r]ubia 'blonde' (examples from Colantoni 2006). Second, Mazzaro (2005) reported that the labio-dental fricative /f/ in Argentinian Spanish undergoes velarization and aspiration (namely, debuccalization) as also found in other Spanish varieties (e.g., New Mexico, Colombia and Venezuela: Quilis 1993). These examples show that onset consonants in some varieties of Spanish undergo quality changes. However, no known changes have been reported for /s/ and /x/ among dialects of Spanish, and we can safely attribute changes in the phonetic properties of /s/ and /x/ to prosodic effects, and not to dialectal differences of the target sounds.

1.4 Research objectives

The main goal of this dissertation is to obtain phonetic details of stress manifestation through which we can better understand how prosody plays a role in segmental variation. This dissertation has three specific objectives. The first objective is to determine the phonetic manifestation of stress that can be identified as stress effects on the constituents of the syllables. Assuming the hierarchical syllable structure, given earlier in (1), we will examine the stress effect on both onset and vowel. This objective will be achieved by answering the following research question, which is further detailed as two sub-questions.

[1]   How does stress affect the phonetic manifestation of syllable in relation to
        the syllable constituency (onset and vowel)?

        a. What are the overall acoustic modifications of syllables influenced by str
        ess?

        b. What are the acoustic changes specific to stressed onset and/or vowel?

The first two questions examine how stressed syllables are acoustically different from their unstressed counterparts. In this examination, the acoustic changes between stressed vs. unstressed onsets and between stressed and unstressed vowels will be compared. The aim is to provide a paradigmatic perspective of stress effects on each syllable constituent. In addition, we will also examine a syntagmatic perspective by examining the stress effect on the onset relative to the effect on the vowel. These two aspects of stress effects will be considered throughout the investigation and will enable us to determine whether the phonetic manifestation of stress in the vowel affects stress effects in the onset. The investigation of stress effects further leads us to the following two questions.

The second objective aims to examine interrelations among the three prosodic factors stress, accent, and position in word, with the following research questions:

[2]   How do stress effects interact with other prosodic effects such as accent and position in word effects?

a. How is the stress effect (lexically motivated prominence) different from accent (phrase-level prominence) and position in word (prosodic boundary strengthening/ weakening) effects?

b. Are there interactions among stress, accent, and position in word effects on C, V, and the syllable as a whole?

The second objective is to determine whether and how stress effects are dependent on or are independent of other prosodic effects (accent and position in word). Stress effects could be similar to those of accents in the sense that both effects would lengthen segmental duration. The stress effect can also be similar to position in word effects, which are known to involve lengthening of initial segments (Fougeron & Keating 1997; Fougeron 2001; Keating et al. 2003; Cho & Keating 2007). Possible differences could be that stress effects are global to the stressed syllable as a domain, while position in word could be local, limiting the lengthening effect to the onset of initial syllable. If so, then the prediction would be that the word-initial stressed onset is the longest when the syllable is accented. While comparing each prosodic effect, we can also

examine degree of lengthening influenced by single or combinatory effects among stress, accent, and position in word. This is addressed as question [2-b].

The third objective is to find evidence of how higher-level rhythmic regularities are involved in temporal modulations by prosodic factors and in possible prosodic interactions that can be constrained by rhythmic regularities assumed in literature. The following question is asked to obtain this objective:

[3]   How is the isochronous rhythm (in Spanish) involved in prosodic manifestation?

More specifically, do prosodic factors interact in such a way that the rhythmic regularity in Spanish can be retained among prominent syllables?

If the isochronous rhythm in Spanish were involved in the prosodic manifestation, we would expect to find a sort of temporal ceiling effects possibly restricted by a temporal regularity among prominent syllables. In addition, the effect of individual prosodic factor may vary depending on how it interacts with other prosodic effects: e.g., reduced stressed effects when accented (and vice versa) or increased effects of individual prosodic factors in the absence of other prosodic factor.

The investigation of the physical aspects of stress/accent/position in word effects will explicate how prosodic effects are phonetically manifested and what the phonetic contributions of each physical manifestation are to making stressed syllables more prominent than their unstressed counterparts. This investigation will also provide us insight into interrelations among prosodic factors, and the systematicity behind the complexity of prosodic effects.

1.5 Organization of chapters

The following chapters are organized as follows. Chapter 2 will describe how speech data were constructed and collected. Overall effects of accentuation in the given speech data will briefly be presented (Section 2.3), in order to justify that accentuation can evidently distinguish the phonetic realization of the target words. Acoustic specifications for measurements will be defined

and two measurements for temporal analysis are used in this study (broad and narrow methods). Chapters 3 and 4 report the results: temporal results in Chapter 3 and spectral and intensity results in Chapter 4. In Chapter 5, we will discuss the acoustic results to answer the research questions we address in the current study. In Chapter 6 we conclude the dissertation with further implications.

Chapter 2 Methodology

2.1 Recording procedure

Among Spanish dialects spoken in Latin America, speakers from Peru and Argentina were recruited. A total of 10 subjects participated in recording: 5 subjects (2 Female and 3 Male) from each of the dialects. The subjects had a length of residence in the United States with an average of 5.3 years, and most came to the US at least after completion of high school levels of education in their home countries.[10] The subjects ranged in age from late 20s to early 40s. All the Peruvian speakers are from Lima, the capital city of Peru, and most of the Argentinian speakers come from Buenos Aires except two who were born in other cities nearby but spent many years in the city for educational purposes.

The recording was carried out on a digital recorder (Marantz PMD 660) at 44.1 kHz sampling rate through a unidirectional condenser microphone (Shure SM48) fixed on a stand and placed at a constant distance of approximately 8 inches. Recording was carried out in a sound-attenuated room in the Department of Linguistics at Stony Brook University. Each recording continued on average for 1.5 hours, including 5 minutes of break time between repetitions. During the first session, one reading of a prepared word list, constructed for this study, was recorded in isolation. It was designed this way so that speakers could become familiar with the words before they used the target words embedded in a frame sentence. After the word reading session, the subjects were asked approximately how many words sounded familiar to them. The average familiarity rate of the target words was reported as 80 ~ 95% by the subjects (132 out of 165 word items). In the second session, the sentences carrying each target word were presented

---

[10] One informant moved to the US with his family before completing high school education at the age of 16. The other has stayed in the US for more than 10 years but completed his college level education in his/her home country. Most of the subjects were either enrolled in the graduate school at Stony Brook or doing research after his/her PhD degree by the time of data collection.

by displaying one sentence per slide on the monitor that was placed 1.5 feet away from the informant. Three repetitions of sentence reading were obtained from each subject.

2.2 Stimuli

CV syllable structure was used to examine stress effects in the domain of the syllable. Each syllable bears one of the two fricatives (/s/ or /x/) in the onset, and one of the five phonemic vowels in the rhyme (/i, u, e, o, a/, all of which are phonemic in Spanish). The target syllables were all open and each of them was embedded in a tri-syllabic word having stress on either the initial or penultimate syllable. For example, the two /se/ syllables in *séquito* and *secádo* are equivalent in terms of PoW (word-initial), but different in terms of stress such that the first one is a stressed syllable. On the other hand, stress in filler words had all possible stress positions, including ultimate stress. Regarding prosodic boundaries, each target syllable is varied in PoW and Position in Intonational Phrase (PoIP). A total of 165 words were constructed; 80 percent of the words served as fillers (the whole wordlist appears in Appendix A-1). The target words that are included in the acoustic measurements are shown in Table 2.1. The first word set (w1~w10) is word-initial, and the other set (w11~w20) is word-medial.

Table 2.1 Word list used in the recording

Word initial					Word medial (non-initial)				
ID	stress	Vowel	/s/	/x/	ID	stress	V	/s/	/x/
w1	stressed	í	sísalo	jícara	w11	stressed	í	risíble	quejíca
w2	unstressed	i	silício	jinéte	w12	unstressed	i	fósiles	lógico
w3	stressed	ú	súbito	júbilo	w13	stressed	ú	desúso	lujúria
w4	unstressed	u	sujéto	jugóso	w14	unstressed	u	cápsula	brújula
w5	stressed	é	séquito	jeparo[1]	w15	stressed	é	cosécha	sujéto
w6	unstressed	e	secádo	jefázo	w16	unstressed	e	díselo	déjelo
w7	stressed	ó	sótano	jódelo	w17	stressed	ó	pasóta	jojoba
w8	unstressed	o	sollózo	jojóba	w18	unstressed	o	kósovo	badajo[2]
w9	stressed	á	sádico	jálalo	w19	stressed	á	biságra	bajáda
w10	unstressed	a	salúdo	jarábe	w20	unstressed	a	pésalo	pájaro

Notes [1] This word (*Jéparo*) was introduced as a possible (non-native) proper name. [2] In this case (*badájo*), the target syllable is final. It is still post-tonic position, which is consistent among word-medial unstressed syllables across the board.

Note that each target syllable differs in its preceding and following segmental contexts, which are known to influence the phonetic realization of the vowel in the target syllables. Therefore, the overall results will be averaged across the items in comparison.

The target words were embedded in carrier sentences that were constructed to accommodate both target words to two positions in intonational phrase (IP): one word-initially and the other one word-medially within a single carrier sentence. For example, target word *reposar* 'v. to rest' appears first in IP initial position and it appears again in the IP medial position in the carrier sentence of (3)-a: *Reposar es un verbo y uso reposar con un adverbio.*

(3) a. [[***Target****$_i$* es un <u>sustantivo,</u>]$_{iP}$ [y uso ***Target****$_j$* con un <u>adjetivo.</u>]]$_{IP}$

'***Target****$_i$* is a noun, and I use ***Target****$_j$* with an adjective.

(Note: the underlined part changes depending on the grammatical category of the target.)

b. [[***Target****$_i$* es el apellido de alguien,]$_{iP}$ [y uso ***Target****$_j$* con su nombre.]]$_{IP}$

'***Target****$_i$* is a last name of someone, and I use ***Target****$_j$* with his/her first name'

c. [[***Target****$_i$* es un imperativo,]$_{iP}$ [y uso ***Target****$_j$* para pedir algo.]]$_{IP}$

'***Target****$_i$* is an imperative form, and I use ***Target****$_j$* to give commands'

The carrier sentence was modified depending on the grammatical category of each word to accommodate each target word within a carrier sentence (3)a-c. The target words are extracted for measurement resulting in a total of 2400: 2 Dialects × 5 subjects × 3 repetitions × 2 Accent × 2 PoW × 2 Stress × 2 Onset fricatives × 5 Vowels.

2.3 Overall descriptions of stress and accentuation effects in the data

This section provides brief characteristics of the speech data obtained, with special attention to how accentuation affects the phonetic realization of syllables. Since we aim to disentangle stress effects from accent effects, we need to establish a context where stress effects are not overlapping with accent effects; e.g., stressed but unaccented versus unstressed and unaccented syllables.

Furthermore, since Spanish has been referred to as a non-deaccenting language in the literature (e.g., Cruttenden 1995), it is important for the current study to justify that our accent conditions correctly provide accented as well as unaccented contexts in comparison.

In the current study, accent variation was implemented by repeating the target word in the carrier sentences: accented target word (new information) vs. unaccented target (given information). Deaccenting of repeated items is obligatory for some varieties of English (Halliday 1967), but it may not be as obligatory in other languages (e.g., Italian and Spanish) (Cruttenden 1995; Wagner 2006). However, a cross-linguistic investigation (in the form of interview) conducted by Cruttenden (2006) shows that Spanish speakers do optionally deaccent given information, though less often than others.[11] The connection between 'givenness' and deletion of accent can also be viewed as a sort of cognitive universal according to Chafe (1976: 31): "The principal linguistic effects of the given-new distinction in English, and perhaps all languages, reduce to the fact that given information is conveyed in a weaker and more attenuated manner than new…"

Even though Spanish has been described as one of the non-deaccenting languages in the descriptive literature (e.g., Wagner 2006), there is evidence for possible deaccentuation in this language (Ortega-Llebaria & Prieto 2006; Ortega-Llebaria 2008). For instance, Ortega-Llebaria (2008) shows that the stress contrast in Spanish disappears in focal contexts, while it is preserved in a post-focal (unaccented) context. In other words, unstressed syllables become as prominent as stressed syllables when focused, resulting in similar phonetic realization for both stressed and unstressed syllables. For the utterance-medial unaccented targets, givenness of the word can either deaccent or weaken the accent effect, leaving possible stress effects in situ. Our data also show that the segmental realization in Spanish can be prosodically distinct depending on

---

[11] One of the examples used in the interview is as follow (Cruttenden 1995): (Question) "Would you like to come to dinner tonight? I'm afraid it's only chicken". (Expected answer) "I don't like chicken". Here, if deaccenting occurs, it is predicted on 'chicken'. The results show that one out of four speakers re-accented (did not delete the accent on 'chicken') in Arabic, German, Greek, Macedonian and Spanish. See also Wagner (2006) for more roles of relative givenness in deaccenting.

accentuation; the phonetic evidence will be provided in the following two sections on pitch realization and degree of consonant lenition.

2.3.1 Pitch contour of accented vs. unaccented words

As a distinct phonological prosodic factor, accent (phrasal-level prominence) in Spanish is mainly cued by pitch perturbation and duration, both of which are also the two cues for stress (Ortega-Llebaria 2006). Ortega-Llebaria describes interactions between pitch and duration in signaling stress and accent, such that flat pitch contour characterizes unaccented (stressed or unstressed) syllables. Her results can be interpreted to mean that pitch contour is one of the most reliable cues that are specific to accent whereas duration can cue both stress and accent. This section will examine first the realization of pitch accent in two target words, in order to identify different characteristics between the two words.

The unaccented condition is implemented by using informational status of the second target word, which is the second occurrence in a given frame sentence; e.g., *Muelelo* (target$_1$: new information) *es un imperativo y uso muelelo* (target$_2$: given information) *para pedir algo* 'Muelelo is an imperative form and (I) use muelelo to give commands'. A representative pitch contour is shown in Figure 2.1. The target word in this figure is *muélelo*, meaning 'Grind it'. The word appears utterance-initially (target$_1$=muélelo$_1$) and then is repeated medially (target$_2$=muélelo$_2$) in the given frame sentence; hence, Target$_2$ becomes given information.

Figure 2.1 Overall pitch contour representing the tonal difference of the target words in a frame sentence

ASF12R1 Target = "muélelo"
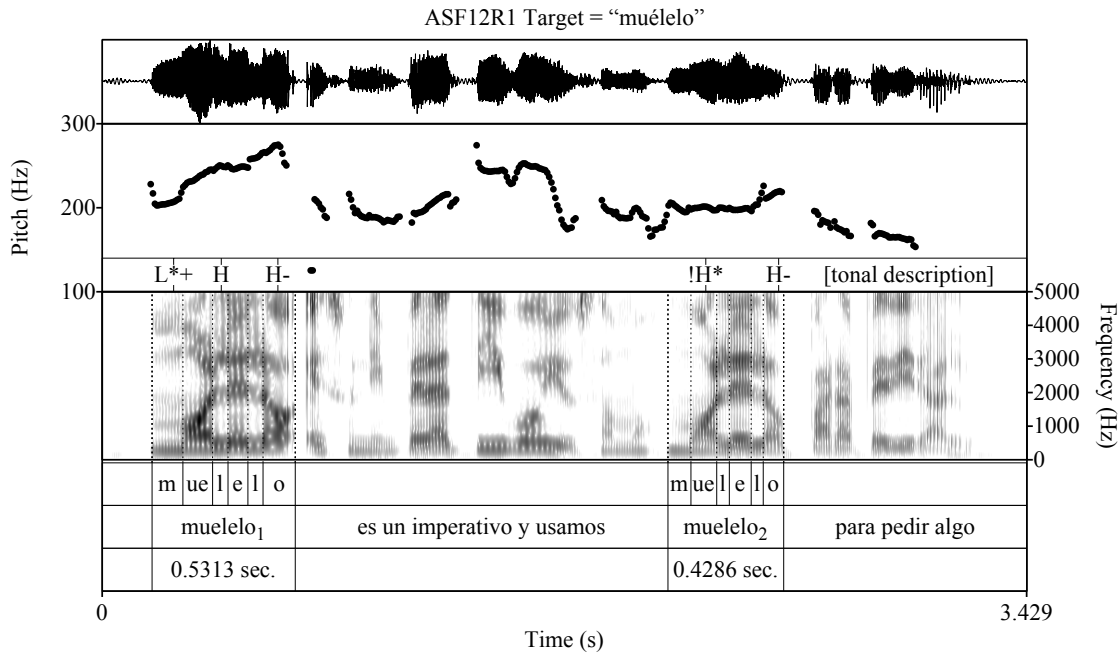
The first target (*muélelo₁*) has dynamic pitch movement over the word; in particular, the initial stressed syllable is associated with a rising pitch target, which can be described as L*+H pitch accent. In contrast, the pitch contour of the second target (*muélelo₂*) is flat over the entire word, and can be described either as lack of pitch accent or as down-stepped H* (!H*) pitch accent (McGory & Díaz-Campos 2000). It has been argued in the literature that pitch is a main correlate of (phrasal) accent (Bolinger 1961; Navarro 1964; Bolinger 1972; Beckman 1986; Ortega-Llebaria 2006). For instance, Ortega-Llebaria (2006) investigated the phonetic correlates of accent and stress in Spanish, and found that both stressed and unstressed syllables displayed a flat intonation contour when unaccented (post-focal position). Thus, the pitch contour in Figure 2.1 can be interpreted to represent either lack of pitch accent or down-stepped pitch accent that is characteristic to the second target word. This flat intonation on the second target word is dominant in the data collected, though occasionally we find clear pitch excursion similar to the initial pitch contour.

31

2.3.2 Degree of stop lenition conditioned by accentuation

Another prominent phonetic difference between the two target words is stop lenition (changes in voicing or spirantization) for voiceless as well as voiced stops. While voiced stop lenition is a well-known allophonic phenomenon in Spanish (e.g., voiced stops /b, d, g/ become [β, ð, ɣ] intervocalically), voiceless stop lenition is more restricted to certain dialects of Spanish (Canary Island and Cuban) and is reported mostly in more recent literature (Guitart 1978; Oftedal 1985; Lavoie 2001; Lewis 2001; Hualde et al. in press). Our data also show that a voiceless stop undergoes lenition (e.g., voicing change) when unstressed. More importantly, this voiceless stop lenition process seems to occur categorically between the two positions: that is, when it occurs, it occurs only in the second target position, and no case is found in the first target. This strongly suggests that the two target words are prosodically distinct and that the second target position is weaker. Figure 2.2 shows two types of phonetic realization of /t/ in the onset of unstressed syllable (*só<u>t</u>ano* 'n. basement'): voiceless stop [t] in (a) and voiced approximant in (b).

Figure 2.2 IP-medial consonant weakening: voiceless stop [t] in unstressed syllables
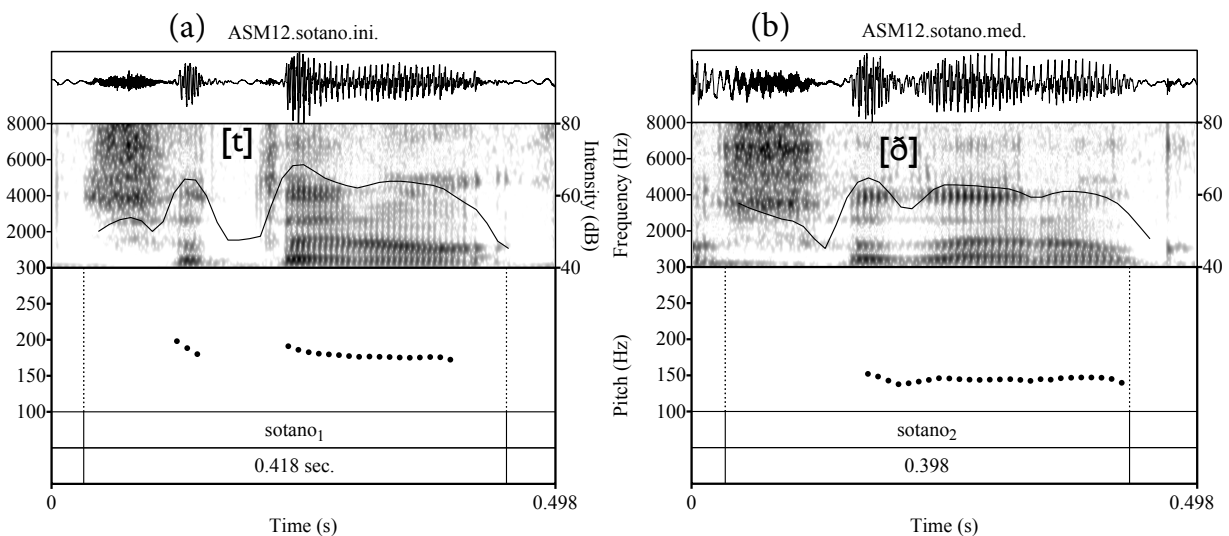
The two targets are taken from one sentence from one of the Argentinian male speakers. The voiceless stop /t/ in (a) is realized with clear stop closure followed by stop burst, but in (b) we see neither the stop closure nor the burst. Another voiceless stop lenition example is shown in Figure

2.3 for one of the Peruvian female speakers. The spectral and pitch information of Target₁ (a) and Target₂ (b) are provided.

Figure 2.3 Stop voicing and weakening: unstressed /d/ and /k/ in *sádico* 'adj. sadistic'
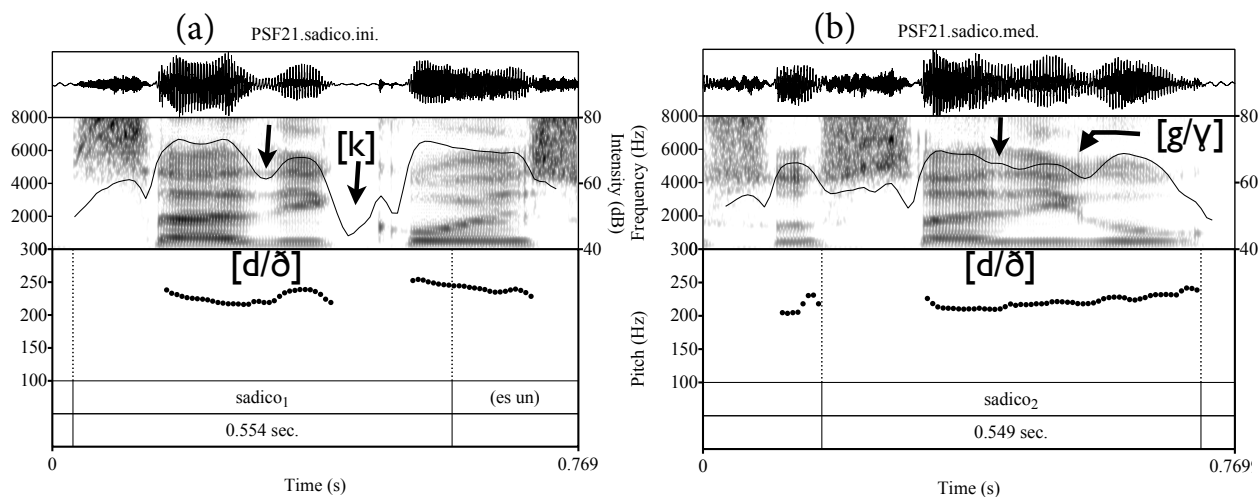
These figures show two types of stop weakening. One is with the intervocalic /d/ and the other with /k/. First, let us compare the intervocalic /d/ in *sádico* 'adj. sadistic' between the two positions. Intervocalic /d/ in Spanish has an allophone [ð] in such a context, so it is not surprising to find the absence of a burst in (a). There seems to be a phonetic trace of closure, as indicated by arrow in the figure (a); also note that the intensity valley appears at this point. This weakening is greater in (b), as indicated by clear formant transition during the consonant and also by the relatively higher intensity level compared to (a). The phonetic behavior of /d/ in these examples indicates that the degree of weakening differs between the two positions and is more extreme in the second position. The second type of consonant weakening is also visible in these figures. Similarly, the intervocalic voiceless velar stop, which is unstressed, weakens to an approximant IP-medially (b), but not IP-initially (a). This is another example of the voiceless stop weakening to a voiced stop (or fricative), as seen in Figure 2.2. Note also relatively flat intonation in (b) compared to (a): e.g., L*+H pitch accent for (a) and !H* or lack of pitch accent for (b). The consonant weakening pattern involved in the second targets also supports the assumption that the two target words are prosodically different in strength. More importantly, the IP medial

targets show phonetic properties that are characteristic of "unaccented" position found in other descriptions of Spanish (Ortega-Llebaria 2006). This fact is taken in this dissertation as a cue for accentuation (phrasal level prominence) such that unaccented words are more vulnerable to a novel weakening process.[12]

Given the phonetic evidence, we conclude that the first target words bear a pitch accent (as a newly introduced item or focus), and the second ones are not accented. For this reason, we will treat the effect shown in the first target as an 'accent' effect.[13] For the purpose of investigating stress effects, we consider mainly IP-medial targets (=unaccented/weakly accented), and we will use the first (accented) targets to compare accent effects to stress effects. The following table summarizes the prosodic characteristics involved in the two targets as well as their segmental properties.

Table 2.2 Linguistic description of target words

Criteria	Target 1	Target 2
Position in utterance?	Initial	Medial
Informational status?	New (and possibly focused)	Given
Pitch accent involved?	L*+H	!H* (or absent)
Degree of voiced stop lenition?	Target 1 < Target 2 (greater)	
Voiceless stop lenition?	Never	Allowed
Accentuation (or if gradient)?	Yes (strong)	No (or weak)
Stress?	Stressed or unstressed	Stressed or unstressed
Position in Word (PoW)?	Word-initial or -medial	Word-initial or -medial

The summarized features will be assumed in the rest of the discussions.

---

[12] Historically, intervocalic voiceless stops in Latin underwent lenition to voiced stops in Spanish; e.g., SAPE:RE (Lt. 'to know') > saber (Sp.), VI:TA (Lt. 'life') > vida (Sp.), AMI:CA (Lt. 'friend, f.') > amiga (Sp.); examples from (Hualde et al. in press). Note that the prosodic context for the change was intervocalic (word-medial) post-tonic position.

[13] It is possible that accent and IP position interact with each other for the cases in our data, where word-initial syllables are in comparison. These syllables are both IP-initial and accented, and thus the initial syllables (or onsets only) may carry both effects. Throughout the dissertation, we will not discuss the two effects separately, and focus on the accent effect that is assumed to be contrastive for the target syllables in comparison.

2.4 Temporal measurement specification

The task of segmentation is always challenging because there is not a simple one-to-one correspondence between the acoustic signal and physiological mechanisms (e.g., Peterson 1955; Stevens 1972, 1989). For the current study, we make use of multiple acoustic landmarks in an attempt to maximize our investigation of acoustic characteristics that seem to be consistently affected by stress and that are possibly accessible to listeners. It is important to define identifiable acoustic boundaries and to do so as reliably as possible. For the purpose of this study, temporal measurements included within-segment temporal characteristics as well as the overall acoustic durations of C, V, and syllable. We will make use of two segmentation methods, broad segmentation and narrow segmentation, each of which will be discussed below.

The 'broad' method takes onset of voicing as the acoustic boundary between consonant and vowel within a syllable. This method follows a traditional type of segmentation; for example, the early work of Peterson & Lehiste (1960) identifies syllable nuclei (vowel) at the onset of voicing, noting that the onset of voicing could be determined relatively accurately.[14] This technical advantage is used in most of the studies on duration implicitly or explicitly.

The broad method assumes that the speech signal is linearly structured; that is, the vowel starts only after the onset C ends. Another assumption is that the acoustic evidence of the vowel depends entirely on glottal excitation, which constitutes only a part of the vowel articulation. Articulatorily speaking, the physical vowel gesture starts much earlier than glottal excitation and/or it is assumed to be coordinated with the gesture of the onset consonant (Browman & Goldstein 1989, 2000; Nam 2004).[15] In effect, there are always continuous movements in the acoustic signal not only between segments but also within a segment. Since those movements are overlapping in the acoustic signal in the temporal dimension, determining exact temporal

[14] Many studies on temporal aspects of speech material have adopted this criterion rather implicitly and rarely explicitly: in particular, it seems to be the default segmentation when unspecified.

[15] For example, the 'In-phase' relation for CV gestures is different from VC phasing that is 'Anti-phase' or sequential (Browman & Goldstein 2000; Nam 2004). The CV phase relation is not reflected in the broad method since voicing onset technically demarcates C and V within the syllable, meaning that V starts with the glottal activation.

boundaries is nontrivial. Nevertheless, we include this method in part because it is so widely used; thus our results can be compared to those of other studies. One supplement to this problem is to also include measurement of acoustically evident transitions. Studies on stop perception have demonstrated the importance of formant changes during the vowel in identifying neighboring consonants. For example, Delattre et al. (1955) demonstrated that the existence of formant transitions (or frequency shifts) near the second-formant loci can serve as cues for the place of articulation of the voiced stops in English, confirming previous results (Cooper et al. 1952; Liberman et al. 1954). Malmberg (1955) further examined the role of transitions in identifying syllable affiliation of stops and found that different inflections in the vowel to consonant transition were perceived as a different syllabic division by the listeners; for example, the perception of [ip-i] or [i-pi] depended on the presence or absence of the consonant induced formant transition carried in the preceding vowel. Such results motivate also including "micro-temporal" dimensions in our analysis. In other words, transitions within and between segments can reflect possible differences that are important in the production and/or perception of speech segments.

The basic idea underlying this narrow measurement is that the details of acoustic cues can be taken advantage of by listeners in identifying specific segments. Consonant cues carried on adjacent vowels are a well-known fact from many studies of both non-sibilant sounds (Liberman et al. 1954; Delattre et al. 1955; Harris 1958) and sibilant sounds (Delattre et al. 1964; Whalen 1981; Bladon et al. 1987; Nittrouer & Studdert-Kennedy 1987; Whalen 1991; Nowak 2006). Even though the fricative noise itself seems to provide reliable cues, many studies have shown that the presence and absence of transitional cues influences perception of specific fricatives. For example, Nowak (2006) examined the role of vowel transitions, and found that transitional cues in the second vowel in VCV words play a role in distinguishing alveopalatal and retroflex fricatives in Polish. In other words, the physical interval that we consider traditionally as a vowel also includes reliable information for the consonant adjacent to the vowel. In addition to the role of the existence of transitions, the temporal phasing of the transition seems to be relevant. Malmberg (1967), for instance, noted that a certain amount of duration is necessary for the listeners to use the transition cues in CVC in identifying the syllabic division of C. This

implies that listeners can interpret the temporal variation of transitional periods. In this regard, the narrow measurement adopted in the current study allows us to gain insight into the effects of stress on some of the more dynamic aspects of speech.

From a methodological point of view, this modular approach may not be a new idea. For example, Hertz (1991) motivates a model that uses phone and transition as explicit units for describing duration patterns of languages, and this phone-and-transition model has been applied to speech synthesis (Hertz 1990, 1991, 1992). For example, she argued that the timing behavior of aspiration and voicing can be captured better by treating transitions as an independent component, than a dependent one as assumed in more conventional models (Klatt 1979; Hertz 1982). Of interest for us is, then, whether the modularized units in the narrow measurement method can capture possible phonetic variation as a function of stress and other prosodic factors such as PoW.

Thus, we also include the 'narrow segmentation' method in our analyses. The narrow measurement marks segment internal temporal points, such as transitions and steady-state portions of consonants and vowels, based on which we attempt to identify sub-phonemic acoustic events. In the following sections, we will define each landmark for the two methods.

2.5 Specifications of broad segmentation

The broad segmentation method marks three boundaries to identify C and V within a syllable: (a) left edge of onset consonant (=syllable left edge), (b) right end of onset consonant, which is identical to beginning of vowel, and (c) right edge of vowel (=syllable right edge). More specifically, the left edge of fricative onset is acoustically identified as the beginning of aperiodic noise and that of vowel is the first glottal pulse (peak) visible on the waveform. For the identification of the right edge of vowel, the final vowel cycle is identified and the closest zero crossing point from the middle of the cycle is taken as the boundary (marked as c in Figure 2.4 below). Complication comes when a target syllable is located word-medially and followed by a voiced stop, which undergoes fricativization in Spanish. This is the case with $SYLL_2$ in '/xoxoba/

(n. jojoba)' (Figure 2.4). In the second syllable (/xo/), identification of the right edge of V_2 is not simple. To resolve this kind of case, the maximum change of formant (max ΔF_2 and/or max ΔF_1) at the acoustically visible juncture of C and V is used as a criterion for the end of V_2.[16] This was determined by visually examining the formant trajectory on the spectrogram, aided by formant tracking over the second half of the vowel. For example, the right boundary of [e] in the figure coincides with the point where F2 changes maximally, moving from the vowel to the following consonant. [17] The following figure exemplifies the acoustic criteria used in the broad measurement.

Figure 2.4 An example of broad segmentation with [$xo_1xó_2ßa$] 'jojoba'
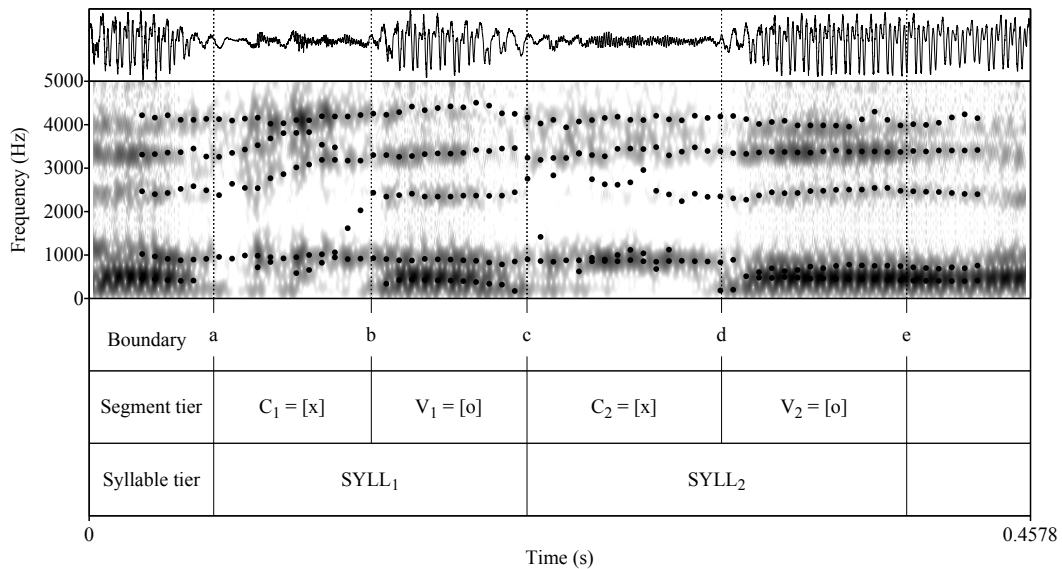
Note: i) Formant settings: Maximum formant was set to 5,5 kHz for 5 formants with 25 ms of window length, and pre-emphasis was set from 50 Hz. ii) Spectrogram settings: Time step of 0.1 (1,000 time steps) was set using Fourier method (Gaussian window). iii) All boundaries are finally marked at zero-crossing point closest to each specific point.

---

[16] For example, a rough landmark between a vowel and the following consonant can be distinguished on the waveform by decreasing amplitude due to consonant constriction.

[17] When the 2 maximum is unidentifiable, F1 and F3 maximum near the boundaries of V served as a subsidiary criterion. This supplementary criterion is consistently applied within a word category.

The target in this case is the initial syllable [xo] in *jojóba*, which is unstressed and word-initial. The beginning of the fricative is identified at the beginning of aperiodic noise, marked as "a" in the figure, and the end of the fricative is at the nearest zero-crossing of the first peak of periodic cycle ('b'), which also marks the beginning of the vowel. This is the point where acoustic energy close to F1 formant starts to accumulate. The end of the vowel is identified in two ways. When a voiceless consonant is following the vowel (e.g., the consonant affiliated to the following syllable), the acoustic beginning of the consonant demarcates the end of the vowel, and this matches a traditional segmentation for vowel ending. On the other hand, if the following consonant is a voiced one, the default criterion specified earlier is used: that is, the F2 maximum changing point is taken as a boundary by examining the formant trajectory in the second half of the vowel and close to the following consonant.

2.6 Specifications of narrow segmentation

For narrow segmentation, five transitional durations and two steady-state parts are identified within the syllable. These include Steady-State periods for C (C-ss) and V (V-ss), and four transitions (T1~T4). The boundaries identified in the broad measurement are included in this narrow segmentation. The additional landmarks adopted in the narrow segmentation identify steady areas for consonant and vowel: these are labeled as Consonant Steady-State (C-ss) and Vowel Steady-State (V-ss). Two acoustic parameters are consulted in determining the boundaries of C-ss and V-ss. The second formant (F2) trajectory is used for V-ss, while intensity is used for C-ss. These additional boundaries are illustrated in Figure 2.5, followed by the acoustic specifications for each boundary.

Figure 2.5 An example of segmentation in the narrow measurement
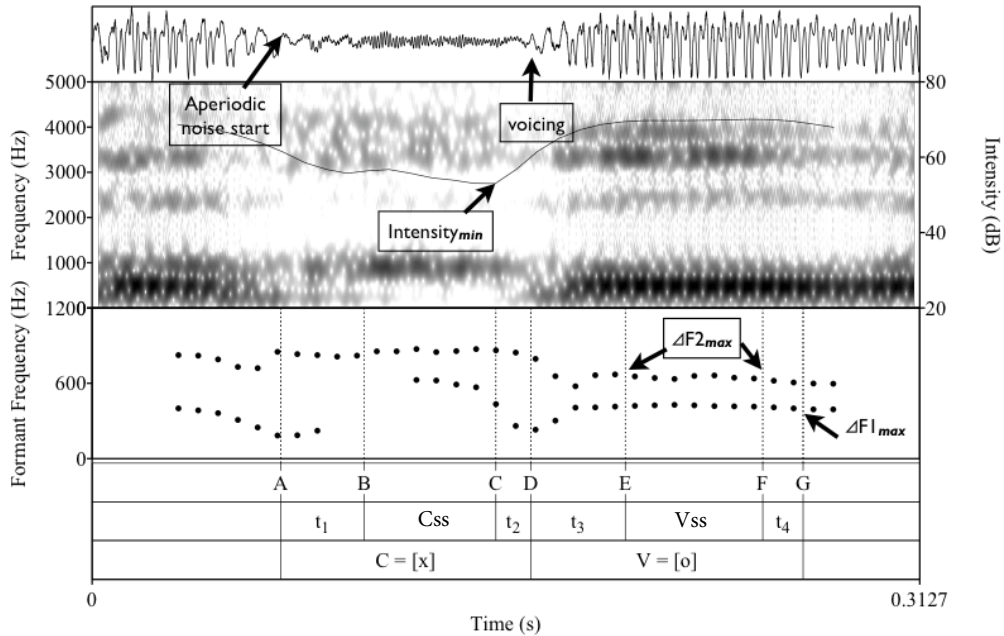
Note: This narrow measurements include Pre-Consonantal transition (T1: A~B), Consonant Steady-State (C-ss: B~C), Post-Consonantal transition (T2: C~D), Pre-Vowel transition (T3: D~E), Vowel Steady-State (V-ss: E~F), and Post-Vowel transition (T4: F~G).

The boundaries labeled "A, D, G" are identical to 'c, d, e' in the broad measurement shown in Figure 2.4. Additional boundaries are 'B, C, E, and F,' each of which marks further transitions and steady-state periods. Point 'B' is taken as the beginning of the main strong fricative noise separating it from the preceding a weak aperiodic noise generation (A~B). An articulatory attribution can be made for this period such that the active articulator is moving to form a narrow channel through which turbulent air is generated resulting in characteristic aperiodic noise for fricatives. Thus, point B signals that a fricative constriction is formed in the mouth.[18] The end of the constriction, marked at point "C", coincides with the intensity valley (minimum). This point is interpreted to mean that the air passing through the fricative constriction becomes weakly turbulent and decreases in energy. Thus, the Consonant Steady-State (C-ss) can be defined

---

[18] In an articulatory study by Cho & Keating (2007), stop consonants showed a significant effect of stress in the articulatory measure of seal duration, even though the overall duration of consonants was not significantly affected by stress. This means that we may observe stress effects (and other prosodic effects) within a segment when measured by the narrow measurement taken in this study.

articulatorily as the period of the narrow constriction that is necessary for fricatives, and acoustically as a time-span of the core aperiodic noise production (Shatz 1954; Stevens 1960). We can make use of C-ss in interpreting how long a fricative constriction is retained.

The second formant (F2) is consulted in determining the landmarks 'E' and 'F'. After the voicing excitation marked as 'D', the vowel formants start to form, in principle reaching the vowel formant values that are characteristic of a specific vowel articulation. Choice of F2 as an acoustic reference is made based on the early research on formant transitions, where the F2 information was recognized as a critical cue for the place of articulation of consonants (Delattre et al. 1955; Liberman et al. 1956; Lehiste & Peterson 1961).[19] The identification of 'E' is made by referring to the F2 trajectory over the first one third of the vowel.[20] The landmark 'E' initiates the Vowel Steady-State (V-ss) and ends at 'F', which is also identified by the F2 change departing from the characteristic F2 value for V-ss. The end of the vowel is the same as defined in the broad measurement: if the following consonant is a voiced one, the F2 maximum changing point is taken as a boundary (point 'F' in Figure 2.5 below) by examining the formant trajectory in the last one third of the vowel and close to the following consonant.

Once each landmark is manually identified through a visual inspection of the spectrogram, a physical boundary is marked at the nearest zero-crossing point to each boundary: technically, a cursor was placed at the mid point between the two formant points (e.g., 6.26 ms distance) on the spectrogram, and then zero-crossing point is queried and selected as a boundary.[21] The maximum F2 turning point could be identified by eye most of the time; in unclear cases, F2 formant value differences were manually calculated based on the formant values extracted in Praat. The acoustic landmarks for each boundary are summarized in Table 2.3.

---

[19] In addition, F2 trajectory is known to reflect coarticulation effects, and can be used to describe coarticulation degree between C and V (Olive et al. 1993: 155).

[20] "One third" of the vowel is consulted in determining each F2 maximum point of change. This means that the vowel steady-state is assumed to exist even though the formant may change continuously over the entire vowel.

[21] The segmentation for the Narrow measurement was all done manually, and repeated 3 times by the author. After this process, labeling was completed for the data randomized within each gender.

Table 2.3 Acoustic landmarks and specifications referred to in the measurements

| Labels | | Phonetic events | Acoustic landmark(s) | |
Broad	Narrow		Left boundary	Right boundary
C	t_1	Constriction starts to form.	A. Aperiodic noise initiation after last complete periodic cycle of a preceding vowel (if present)	B. High frequency noise starts to form (characteristic frequency noise band is visible on the spectrogram)
	C-ss	Consonant steady state; acoustic constriction target is reached.	"B" as defined.	C. Intensity minimum syllable internally
	t_2	Constriction begins to release and intensity control switching from C constriction to V opening.	"C" as defined.	D. First glottal pulse
V	t_3	Voicing starts.	"D" as defined.	E. F2 maximum turning point for the one third of the vowel. Afterward F2 (& F1) is relatively stable.
	V-ss	Vowel steady state; Acoustic vowel target(s) reached and sustained.	"E" as defined.	F. F2 maximum turning point for the last one third of the vowel to consonant transition.
	t_4	Acoustic vowel release.	"F" as defined.	G. Maximum point of spectral change of F1 (or F2) for the last one third of the vowel to consonant transition.

Note: The left boundaries (b~f) are identical to the right boundaries (b~f).

The acoustic landmarks are exhaustively applied and therefore the left edge of the next unit coincides with the right edge of a preceding unit. The only exceptions are the syllable initial and final boundaries. For these, both waveform and wide band spectrogram were consulted in the segmentation process. The spectrogram window was 5ms (wide-band; around 260 Hz bandwidth) with a Gaussian window shape. The maximum formant value was set at 5 kHz for male and 5.5kHz for female in calculating the first 5 formants (10 ms of time step). The sampling rate of the speech file was 4.4 kHz.

2.7 Spectral and intensity measurements

We examined characteristic changes in spectral parameters for the fricatives and the vowels as a function of stress. Spectral moments and spectral balance are measured for the onset fricatives and the vowels, respectively.

Firstly, stress effects on spectral characteristics of the onset fricatives are examined by comparing the changes in the four spectral moments: Center of Gravity (COG=mean frequency; M1), Variance (M2), Skewness (M3), and Kurtosis (M4) of the fricative spectra. For the first moment measure of the fricative spectra, a higher COG is expected for fricatives with a shorter anterior resonating cavity: thus, COG for /x/ should be considerably lower than that of /s/.[22] COG tends to increase as a fricative is articulated farther front in the mouth (narrower cross-sectional area in the front-cavity) and to decrease as a fricative constriction is formed in the further back (more open cross-sectional area) (Stevens 1998). For fricatives proper, the vertical size and shape of a constriction correlate with the major frequency distribution, together with the velocity involved in the articulation of each fricative. The higher the COG values appear in the spectrum, the smaller the front cavity formed (Stevens 1999). In addition, COG is influenced by the size of the cross-sectional area of a constriction in the front-cavity as well as the aerodynamic condition involved. That is, the generation of frication noise requires an increased intraoral pressure, and this aerodynamic condition has to be met for the size of the constriction (Shadle, 1990; Stevens, 1997). In the literature, spectral COG is used in describing different places of articulation for fricatives (Jassem 1979; Cho & Keating 2009). For instance, Cho & Keating (2009) found prosodic effects as well, such that the COG of stop burst is influenced by stress as well as accent. Kim (2006) showed that English fricative /s/ was affected by stress and other prosodic factors in a number of phonetic properties including COG. These cues in turn influenced Korean listeners' perception of English [s] either close to a lenis fricative or as a tense

---

[22] Spectral outcome of fricative production depends partly on the size and shape of the oral cavity in the vicinity of a fricative constriction, through which turbulence noise is generated and the source noise at the glottis is filtered (Fant 1960; Stevens 1998).

fricative, two of which are phonemically contrastive in their native language. These findings suggest that stress effects (and prosodic effects in general) can be reflected in the spectral COG.

Besides the COG measurement, variance (M2: the second moment) describes how the overall energy distribution is spread from COG in the frequency domain. The third and the fourth moments can be used to describe whether more energy is tailed toward left (lower frequencies) or right (higher frequencies) from the mean frequency, and how heavy the tails in the spectral energy distribution would be over the frequency domain. It is assumed in the current study that COG values are sensitive to the fricative constriction following the aforementioned literature, and that it can give us a way to look at how articulation is influenced by stress for each onset fricative. We are particularly interested in determining whether stress effects are involved in modifying spectral characteristics within each fricative category. Our prediction is that stressed fricatives would show relatively higher COG values compared to the unstressed counterparts. One empirical question is to which degree the two dialects show stress effects on the spectral characteristics and we will keep this question in mind when we discuss the results in Chapter 4.

For the acoustic analysis of COG, a spectral slice of each C-ss interval is extracted using a Hanning window after downsampling to 22,050 Hz. An LPC smoothed spectrum was used to check the output COG value.[23] A Praat script was employed for the spectral measurements. For the vowels, energy distributions are examined as a measure of spectral balance above 0.5 kHz. The higher energy is measured over the four filtered bands that are contiguous: B1(0~0.5kHz), B2(0.5~1kHz), B3(1~2kHz), and B4(2~4kHz).[24] The investigation of spectral distributions in the vowels enables us to define stress effects as more dynamic changes rather than static characteristics that refer to a mean value within a phonetic parameter. It is generally understood that overall intensity correlates with stress (Lehiste 1970; Beckman 1986), although the perceptual role of intensity is either absent or, if found, weaker than other parameters in the identification of stress (Fry 1955, 1958; Beckman 1986; Llisterri et al. 2003). As Lehiste (1970) pointed out, although stressed syllables are generally longer and also louder than unstressed syllables, the lack

---

[23] LPC smoothing was set with 6 poles over 11,025 Hz, and no pre-emphasis was set (Ltas, 1-to-1).

[24] For these band specifications, see Sluijter & van Heuven (1996b).

of a direct relationship between stress and intensity seems to be due to the fact that intensity depends in part on other physiological factors such as the articulatory configuration in the vocal tract and subglottal pressure. This explains why intensity has been reported as a weak cue and its role in the perception has often discussed in combination with other cues such as pitch. According to Sluijer & van Heuven's (1996) study, intensity as one of the stress correlates can be better understood by examining the energy changes over frequencies, namely spectral balance, rather than overall intensity levels. They found that only the intensity at higher frequencies changed as a function of stress, and not the lower energy (below 0.5 kHz). Spectral balance as a stress correlate has been confirmed by other studies: American and British English (Turk & Sawush 1997; Turk & White 1999), Catalan (Astruc & Prieto 2006), and Spanish (Ortega-Llebaria & Prieto 2006).

Stress effects found in spectral balance in vowels can be understood to mean that articulatory effort (e.g., extra changes/movements of articulators compared to normal speaking articulation) produces higher energy. For instance, Glave & Rietveld (1975) examined the effects of varying effort on the spectral intensity distribution, and found that there was greater intensity in the higher frequency region (above 0.5 kHz) of vowels when the speakers put greater vocal effort in their speech production. Under the assumption that stress realization correlates with some form of vocal effort, we would expect to see the stress effect in our data. Our main interest is to determine whether the differences in spectral balance could be a reliable cue (even when unaccented), and whether the stress effect in the spectral balance interacts with other prosodic factors, such as accent and PoW.

Finally, vowel formant changes as a function of stress are also investigated. The influence of stress on vowel quality is abundantly evident in the grammar of spoken languages. For instance, many stress system languages show some degree of restriction on phonemic vowel contrasts depending on the presence or absence of stress. In general, reduced vowels can be found in unstressed syllables, while full vowels appear in stressed syllables and/or are only selectively allowed in unstressed syllables (Partridge 1950). In addition, a number of studies have explored the acoustic correlates of stress and concluded that vowel quality is affected by stress as well as other phonetic factors (Bolinger 1958; Lieberman 1960; Beckman 1986; Sluijter & van Heuven

1996b). For example, stressed syllables are associated with more peripheral vowels than unstressed syllables, which in turn tend to be associated with less peripheral and centralized/neutralized vowels (Jones 1956; Tiffany 1959; Lindblom 1963; Delattre 1969; Gay 1978; Fourakis 1991; Cho 2005).

Stress effects on vowel quality in Spanish are described as rare (Hualde 2005). This is evident from the fact that studies examining stress effects in Spanish do not consider vowel quality as a parameter. Although stress effects on vowel quality in Spanish may not be as evident as in other languages such as English, Dutch, and Swedish, we find descriptions suggesting that vowel quality in Spanish is also influenced by stress. For example, it has been reported that speakers tend to reduce the F1-F2 distance in unstressed [e] (Ortega-Llebaria & Prieto 2005; Ortega-Llebaria 2008). The effect, however, decreases in post-focal sentences and disappears in focal contexts. This means that vowel quality in Spanish is sensitive to prosodic factors such as accent and stress. In addition, we will see in Chapter 3 that vowel durations were significantly different between stressed and unstressed conditions. If vowel reduction depends entirely or partly on the duration of vowel (e.g., Lindblom, 1963; Sluijter & van Heuven 1996), we would expect stress and/or accent effects in the vowel formant. This chapter examines prosodic effects on the vowel quality as measured by the first two formants.

For the vowel formant analysis, the time step was set to 0.01 with a 25 ms window length. A Praat script was used to calculate the formants. F1 and F2 are calculated using 'To formant (burg)' and the tracker, whose parameters were set for each gender.[25] The script allows the user to manually check the formant values referring to Ltas (Long-term average spectrum) and the LPC spectrum (Linear predictive coding; 'autocorrelation' algorithm).

---

[25] Formant tracking (3 paths) is set at 500, 1485, and 2475 for male, and at 550, 1650, and 2750 for female (suggested averaged formant values in Praat manual; http://www.fon.hum.uva.nl/Praat/manual/Formant__ Track___.html). The script is slightly modified for practical purposes of this study from the original version created by Dr. Bert Remijsen (http://www.ling.ed.ac.uk/~bert/Praatscripts.html#msr&check_f1f2_indiv_ interv.psc).

Overall energy is also examined for both fricatives and vowels, as measured by Root Mean Square (RMS, in Pascal) of the speech signal.[26] Since each syllable is extracted from different speech recordings, we use ratios of the consonant's RMS in relation to that of the vowel within the same utterance to normalize variation potentially induced by loudness in a given context. The RMS values are obtained using a Praat script, for the intervals C-ss and V-ss.[27] When the ratio of C-ss to V-ss increases as a function of stress, the interpretation is that the consonant energy is increased relatively more than for the vowel. Thus, we can see how stress affects the intensity relation between onset fricative and vowel in a syntagmatic way.

Having the results from both the temporal and spectral dimensions, we will be able to better understand stress effects on segments. In addition, the integration of the results will enable us to understand how stress could be implemented in a language (dialects), and also how multiple parameters interact in stress manifestation. Spectral and intensity characteristics of stress effects will be used as a secondary reference to further investigate how the quality of segments are interrelated to the temporal variation induced by stress.

---

[26] RMS (=mean amplitude) in Pascal is the fundamental measure of sound pressure, based on which sound pressure level (dB) is converted to logarithmic scale: 1 Pa = 93.98 dB. For example, a speech sound is in general 0.2 Pa (close to 60 dB). The RMS measurement for our purpose is to calculate RMS C/V ratio within the syllable, and we use directly RMS values to calculate the ratio result.

[27] Root-mean-square value is defined in Praat as follows: $\sqrt{(1/(t2-t1)\int_{t1}^{t2} x^2(t)dt)}$, where x(t) is the amplitude for the sound and t1 and t2 is the time domain

Chapter 3 Temporal modulation by stress

Cross-linguistically, stressed syllables are generally longer in duration when compared to unstressed syllables. In this chapter, we discuss duration results focusing on whether stress-induced temporal modulation varies at the level of dialects, and how the syllable constituents (onset and rhyme) behave as stress changes. The duration results are reported based on the two methods defined in Chapter 2: the broad and the narrow measurements. The syllable is assumed to be the domain of stress realization, and thus is expected to undergo specific temporal modulations that are motivated by stress manifestation. An additional factor to consider is the degree to which stressed and unstressed syllables differ in duration, which may vary from one language to another. For example, unstressed syllables in English may undergo durational reduction in most cases, and can even be reduced to a syllabic coda without a nucleus vowel. This is unlikely to be found in Spanish, though unstressed syllables are generally shorter than their stressed counterparts (Quilis & Esgueva 1983; Hualde 2005). Conversely, stress effects will be identified exclusively from the acoustic characteristics found in the stressed syllables: in other words, if the phonetic properties are distinct from those found in unstressed syllables, and unless there are known contributing factors other than stress.

This chapter is organized as follows: section 3.1 reports the duration results based on the two measurements (Broad and Narrow). Stress effects are investigated first for the syllables as a whole, and then for the onset fricatives and the vowels, respectively. Section 3.2 examines PoW effects, and discusses how the stress effect interacts with the PoW effect. Finally, section 3.3 summarizes the duration results with discussions of the findings.[28]

---

[28] One conditional factor present in our data is gender. For each dialect, we have 3 male and 2 female speakers. When gender difference seems to be a main factor in any pattern, the results will be separately examined by gender.

3.1 Broad measurement results

First we use the Broad duration measurement method to examine degree to which durational changes of /s/V and /x/V syllables vary depending on changes in stress and accent. Stressed syllables are expected to be longer than their unstressed counterparts, and this is what we find in Table 3.1. A total of 2400 tokens are included from the accented and unaccented syllables in this analysis.

Figure 3.1 Stress and accent effects on syllable duration
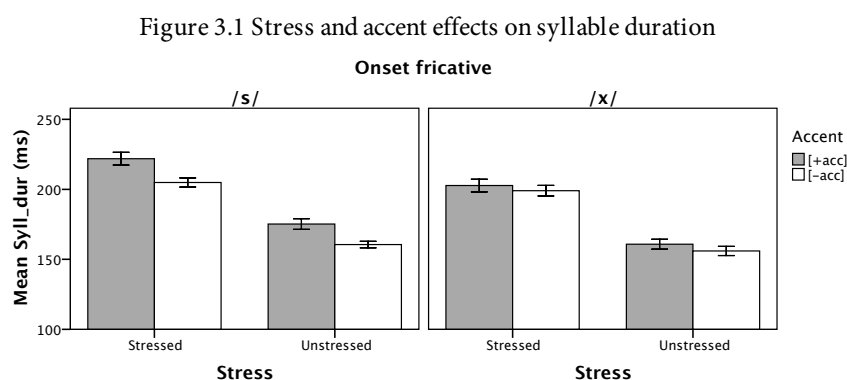
Table 3.1 Averaged syllable duration of /s/V and /x/V sequences

Dialect	Accent condition	Stress condition	(a) /s/V syllables Duration (SD) in ms	(b) /x/V syllables Duration (SD) in ms
Argentinian Sp. (N=1200)	Accented	Stressed	217.7 (40.4)	199.2 (34.8)
		Unstressed	171.9 (27.9)	156.4 (26.2)
	Unaccented	Stressed	203.1 (26.4)	199.1 (30.4)
		Unstressed	157.9 (18.9)	154.8 (29.2)
Peruvian Sp. (N=1200)	Accented	Stressed	225.8 (33.8)	206.0 (39.1)
		Unstressed	178.3 (33.6)	164.9 (30.7)
	Unaccented	Stressed	206.5 (27.2)	199.1 (31.8)
		Unstressed	162.7 (20.2)	156.9 (25.0)

Let us first examine the /s/V syllables. The mean duration of syllables is longer for the stressed condition than the unstressed one, and both the stressed and unstressed syllables are longer when they are accented. This means that accent influences duration regardless of stress for the /s/V

49

syllables. This effect is weaker in the /x/V syllables. Although the syllables tend to be longer in the accented condition, the difference is not statistically significant ($p>0.05$). Thus, we can conclude that accent effect is present but differs depending on the onset type. Repeated Measures ANOVAs show that Accent, as a main effect, is not significant ($p<0.05$). There is an interaction between accent and C_type [$F_{(1,9)}$=11.33, p=0.008]. This interaction indicates that the accent effect is significant for /s/V [$F_{(1,9)}$=6.53, p=0.03] but not for /x/V ($p>0.05$).

The stress effect is consistent regardless of Accent and onset type. The averaged durations of syllables are longer in the stressed syllables by 42~45 ms. The differences in duration are almost identical between the two onset fricatives (/s/ and /x/), and very similar between the two dialects (28% for Argentinian Spanish and 26% for Peruvian Spanish). According to Repeated Measures ANOVA, there is a main effect of stress on duration and no interaction between Accent and Stress: [STRESS ($F_{(1, 9)}$=186.72, $p<.001$); STRESS ACCENT ($F_{(1, 9)}$=0.047, p=0.833].[29]

The durational results of syllables are comparable to previous studies; for example, the averaged syllable durations in Gutiérrez-Díez (2001) varied from 148 ~ 211 ms depending on the number of syllables of the words (1~7).[30] In another study by Engstrand & Krull (2002), the durations of CV syllables were recorded as 198 ms when stressed and 134 ms when unstressed in the reading speaking condition. They also found that syllable durations tend to decrease as speaking conditions become more casual: the average durations of syllables decreased to 170 ms (stressed) and 144 (unstressed) in the unscripted condition.[31] Thus, it makes sense that the mean

---

[29] The syllable durations are generally longer when the onset is /s/ pooled across dialects and stress conditions: C_type ($F_{(1,1119)}$ =7.01, $p<.008$) according to One-way ANOVA.

[30] The syllable durations in this study show that syllables from mono-syllabic words are the longest (also those from the 5 and 6 syllabled words). In other cases, the durations were similar for the syllables embedded in the 2-4, and 6 syllabled words.

[31] The types of segments were not specified but presumably the mean durations are from various segments included in the text they used. For the unscripted speech, they described, "it consists of lively monologues produced by one male native speaker (Argentinian)."

durations obtained in the current study are closer to their result from the reading condition: 204.8 ms for the stressed and 159.3 ms for the unstressed syllables.

Durational patterns can be affected by many factors, such as PoW, number of syllables in a word, informational load of a word, and presence/absence of boundary nearby (Delattre 1966; Umeda 1977; Hoequist 1983b; Lavoie 2001; Engstrand & Krull 2002). Amongst the factors, stress has most often been recognized as important in determining the durational realization of consonants/syllables (Hoequist 1983b; Lavoie 2001). For example, Hoequist (1983) emphasized the influence of stress as a main factor that determines language-specific durational differences between syllables.

The syllable durations are now decomposed into the durations for C and V to examine whether the accent effect differs between the syllable constituents. The accent effect on the durations of the two fricatives (a) and the vowels (b) are compared in Figure 3.2. The accent effect is further separated by stress condition.

Figure 3.2 Durations of C and V as a function of accent

(a) C duration
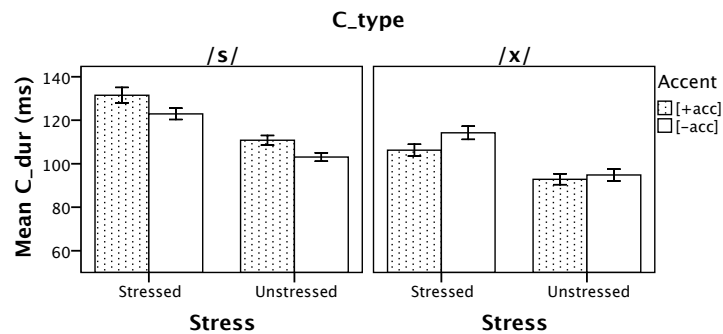
(b) V duration
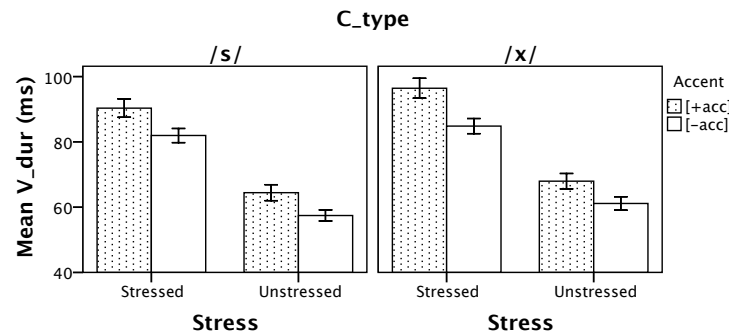
The accent effect on the duration of C is different between /s/ and /x/. The duration of /s/ is longer when accented, regardless of stress. The duration of /x/ is the opposite: unaccented /x/ is longer when stressed. No effect is found for the unstressed /x/. On the other hand, the accent effect on the vowel duration is consistent regardless of the onset fricative type: vowels are longer when accented, and this effect is independent of the stress effect. In other words, stressed vowels are the longest when accented, and unstressed vowels are the shortest when unaccented. Repeated Measures ANOVAs confirm that there is an interaction between Accent and C_type [$F_{(1,9)}$=24.28, p=0.001]. The accent effect on the vowel duration is significant [$F_{(1,9)}$=7.42, p=0.023], but not on the consonant duration [$F_{(1,9)}$=0.026, p=0.88]. Within the C_type category, the accent effect is near significant for /s/ [$F_{(1,9)}$=5.08, p=0.051] and not for /x/ [$F_{(1,9)}$=4.0, p=0.76]. From the results, we can conclude that the accent effect is only consistent for the vowel such that vowels are longer when accented regardless of stress condition. For the consonant, the accent effect is inconsistent, and depends on the onset type. Our results show the accent effect only for /s/.

Now we examine the stress effect. For the purpose of the current study, the IP-medial target syllables are considered since this is the place where we assume the stress effect is not overlapping with the accent effect. The averaged durations of C, V, and syllables are summarized: /s/V syllables in Table 3.2 and /x/V syllables in Table 3.3. The mean durations of consonants /s/ and /x/ and the vowels appear separately from the total syllable duration.

Table 3.2 /s/V syllables: averaged durations of syllable constituents by stress

| | | Duration in ms (SD) | | |
		C = /s/	Vowel	Syllable
Argentinian Sp.	Stressed	119.27(18.89)	83.79 (18.82)	203.06 (27.78)
	Unstressed	101.85 (15.43)	56.08 (15.41)	157.93 (18.89)
Peruvian Sp.	Stressed	126.36 (23.41)	80.11 (16.54)	206.47 (27.16)
	Unstressed	104.13 (16.64)	58.58 (12.05)	162.71(20.17)

Table 3.3 /x/V syllables: averaged durations of syllable constituents by stress

		Duration in ms (SD)		
		C = /x/	Vowel	Syllable
Argentinian Sp.	Stressed	110.05 (23.50)	89.01 (19.92)	199.06 (30.40)
	Unstressed	94.47 (22.12)	60.35 (18.24)	154.81 (29.22)
Peruvian Sp.	Stressed	118.10 (25.46)	80.96 (17.61)	199.06 (31.83)
	Unstressed	95.07 (22.81)	61.86 (15.53)	156.93 (25.04)

First, note that both onset fricatives and vowels are longer when stressed: Δdur (str-unstr) = 17 ~ 22 ms for /s/, and 21~27 ms for the vowel. Repeated measures ANOVA shows that STRESS as a main factor is significant for the consonant duration ($F_{(1,7)}$=67.8, $p<$.001), and for the vowel duration ($F_{(1,7)}$=210.8, $p<$.001). The results confirm that both C and V are active in the temporal modulation of syllables when stress is involved, and the vowel tends to be more influenced by stress at least in the temporal dimension.

For representational purposes and further discussion, the stress effect on the syllable durations is visually summarized in (a) (/s/V syllables) and (b) (/x/V syllables) in Figure 3.3.

Figure 3.3 Durational composition within the syllable: stress effects on /s/V and /x/V syllables

(a) /s/V syllable duration by stress

(b) /x/V syllable duration by stress
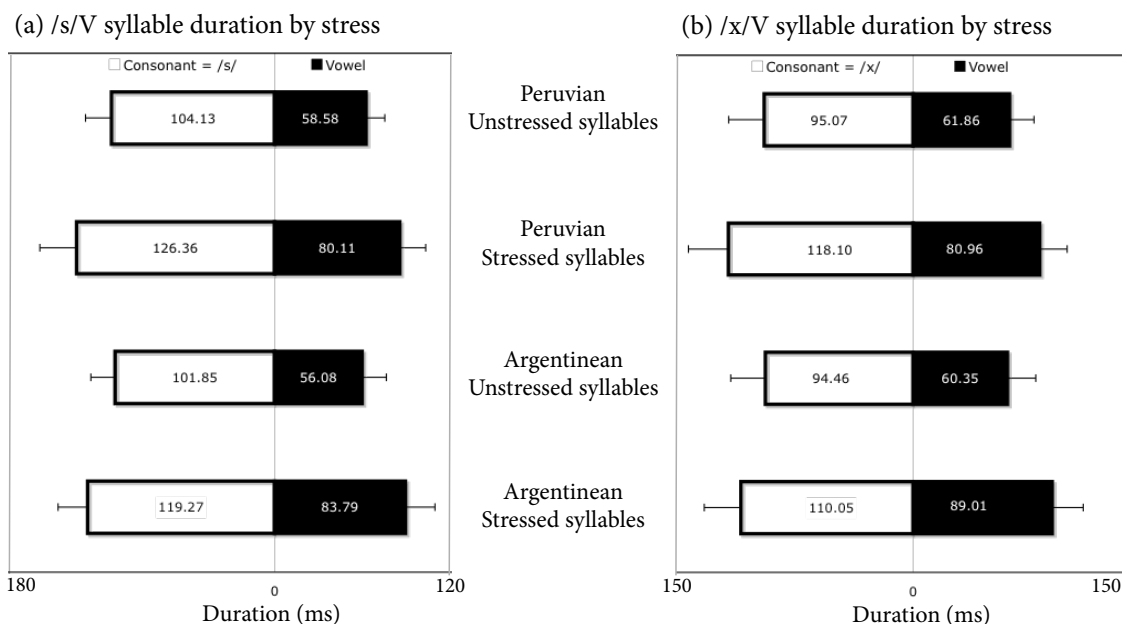
When the durational changes are calculated proportionally from unstressed to stressed syllables, we see that the vowel duration is more sensitive to stress than the consonant, and these durational patterns are consistent across the two types of syllable groups and the two dialects.[32] For example, Peruvian and Argentinian Spanish show 22.4 % and 17% increases in the onsets, and 39% and 49% in the vowels, respectively. Between the syllable constituents, the durational increase of onset is greater in Peruvian than in Argentinian Sp., and the opposite trend is found in the vowel: [Dialect as a factor on C duration: $F_{(1, 599)}$=24.16, p<0.001; on V duration: $F_{(1, 599)}$=2.65, p=1.04]. According to Pearson's Product-Moment Correlation analysis, there is a significant correlation between DIALECT and C-duration ($r_{(1198)}$=.113, p<.01; 2-tailed), but not between DIALECT and V-duration ($r_{(1198)}$=.007, p=.8). This may be interpreted to mean that longer fricatives tend to be affiliated with Peruvian rather than with Argentinian.[33] The mean differences of C, V, and syllable durations by gender are slightly longer for female speakers than male speakers, but the difference is not statistically significant.

The durational ratios of syllables in our data are provided in Table 3.4. The averaged ratio of the durations is 1.3, pooled across conditions, meaning that the stressed syllables are 30 percent longer than the unstressed counterparts in duration. This duration ratio by stress is the same regardless of the onset fricative type.

---

[32] As we compare the two figures, the varying degree of sensitivity to stress seems to pattern similarly for /s/V and /x/V syllables; there was no significant interactions between the two types of syllables as confirmed by two-way ANOVA analyses [C-type × Stress = ($F_{(1,9)}$=.209, p=.66) for C-duration; $F_{(1,9)}$=.854, p=.38 for V-duration; and ($F_{(1,9)}$=1.15, p=.31) for syllable-duration].

[33] There are significant correlations between C-duration and V-duration ($r_{(1198)}$=.220, p<.01**; 2-tailed) across the stress conditions. This correlation is interpretable referring to the durational behavior represented in Figure 3.3 and Table 3.4. For example, the main effect of stress is to increase both onset and vowel in duration, and therefore a longer vowel predicts a longer consonant.

Table 3.4 Means of the duration ratio between stressed and unstressed syllables

PoW	Dialect	Mean (SD)		
		/s/	/x/	When pooled across
Word-initial	Argentinian Sp.	1.31 (.16)	1.25 (.25)	1.28 (.21)
	Peruvian Sp.	1.25 (.18)	1.27 (.23)	1.26 (.20)
	Total (N=150)	1.28 (.17)	1.26 (.24)	1.27 (.21)
Word-medial	Argentinian Sp.	1.28 (.22)	1.38 (.24)	1.33 (.23)
	Peruvian Sp.	1.37 (.28)	1.33 (.22)	1.35 (.25)
	Total (N=150)	1.33 (.26)	1.35 (.23)	1.34 (.24)
Total	Argentinian Sp.	1.30 (.19)	1.31 (.25)	1.31 (.22)
	Peruvian Sp.	1.31 (.24)	1.30 (.23)	1.31 (.23)
	Total (N=600)	1.31 (.22)	1.31 (.24)	1.31 (.23)

A ratio analysis for durational behaviors under stress has been explored in other studies, as a way to normalize raw data: 1.23 in Delattre (1966), 1.16 in Olsen (1972) from a Mexican Spanish speaker, 1.39 from a reading passage from a Castilian speaker in Gili Gaya (1940: reference from Gutiérrez-Díez, 2001). Although a direct comparison of different numbers would not be available since each study differs in the dialects of the subjects and recording material, it helps us to assure that the overall stress effects on the syllable durations in our results (1.3 approximately) are quite comparable to what has been described as stress effects in the literature. We can also see how stress effects on the overall durations interact with other prosodic factors (e.g., PoW), which will be examined later in this chapter.

The durational behavior of C and V as a function of stress was examined in Ortega-Llebaria (2006) for five Spanish speakers.[34] The results were discussed in terms of C/V ratio of the syllables taken from the pseudo word *a-ga-ga-gó*, which was either located in a normal declarative sentence or in a parenthetic phrase. The former context places the target word into a pitch accent position of declarative sentence while the latter expects no overlap of pitch accent with stress. She found that the C/V ratio is higher in the stressed syllables regardless of the presence/absence of pitch accent, and that the ratio of stressed syllables increases from the absence to the presence of

---

[34] Her study recorded speech from 2 male and 3 female (four of them from Spain and one from Mexico).

pitch accent. The result suggests that the effect of stress on duration is independent from that of pitch accent in Spanish.

The C/V ratios were calculated from the current study as fricative duration over the onset over vowel duration. Since the durations of fricatives are found to be longer than those of vowels, C/V ratios will be greater than 1 in most of the cases, and a decrease in C/V ratio would indicate a greater proportional increase of V duration relative to C duration within syllables. The following two figures (Figure 3.4 a-b) show the averaged C/V ratios within the stressed and the unstressed syllables. The results are separated by onset fricative /s/ and /x/ (left vs. right) and by dialect (top vs. bottom). Accent effect is compared in Figure 3.4.

Figure 3.4 C/V ratio and consistent stress effect independent of accent effects

(a) Accented
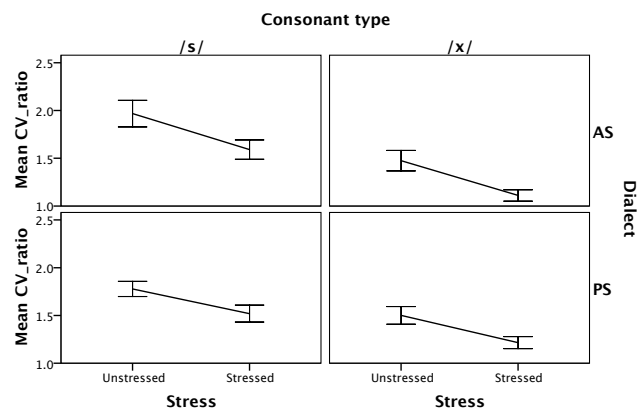
(b) Unaccented
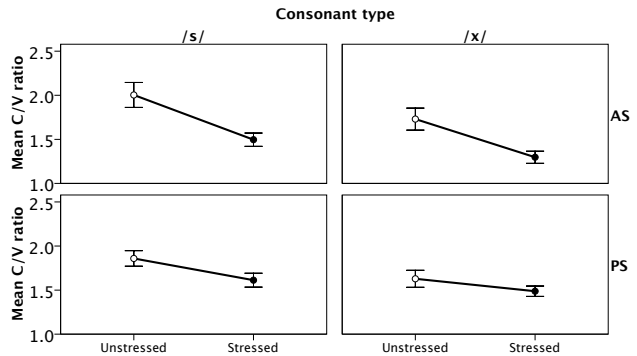
Figure 3.4 shows three characteristic trends in the changes of the C/V ratio. First, the C/V ratio decreases from the unstressed syllables to the stressed ones, and the difference in the ratio

between the stress groups is statistically significant: Repeated Measures ANOVA, [$F_{(1, 9)}$= 36.2, p<.001]. Secondly, this effect is consistent regardless of the accent condition (p>0.5), meaning that the overall durational changes are induced mainly by stress. Finally, the extent to which stress affects each syllable constituent differs between the two dialects as reflected in the C/V ratio values [$F_{(1, 9)}$= 6.7, p<.034. The C/V ratio changes by stress are greater in the Argentinian syllables than the Peruvian ones. There is no main effect of the onset type on the result of the C/V ratio (p>.68).

The consistency of the stress effect on the onset fricatives is confirmed further by examining different vowel groups that are known to differ intrinsically in duration. That is, if the durational change induced by the stress effect is consistent, then the C/V ratio will also change as a function of the durational differences in vowels. While the durations of the consonants vary consistently as a function of stress, we predict that the ratios would be higher for high vowels and lower for lower vowels, if vowel durations are intrinsically shorter for the high vowels. This prediction was borne out, as seen in Figure 3.5.

Figure 3.5 C/V ratio by stress in relation to the vowel height
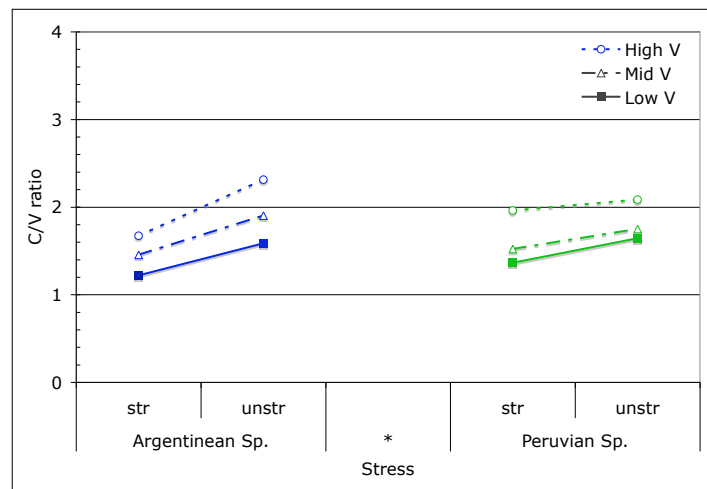
The C/V ratio changes by stress in Figure 3.5 can be interpreted to mean that stress-induced durational changes are not symmetric between consonant and vowel: There is a greater effect on the vowel in duration. As predicted, the C/V ratio for the high vowels is the highest, followed by

mid and low vowels (HEIGHT: $F_{(1,9)}$=109.17, p<.001). This pattern is consistent between the two dialects. What differs between the two dialects is the extent to which the C/V ratios change from stressed to unstressed syllables; the difference is greater in Argentinian than in Peruvian Spanish. The greater the differences in the C/V ratio between stressed and unstressed syllables are, the more difference there is in stress effects on different syllable constituents. A series of Repeated Measures ANOVAs also confirms that the differences in C/V ratio between the two stress conditions are statistically significant (p<.05). Overall, the stress effect on C/V ratio differs between the two dialects [STRESS × DIALECT: $F_{(1,9)}$ =6.4, p<.05].

Although the ratio values in our study cannot directly be compared to, e.g., Ortega-Llebaria's (2006) results due to different onset segments, we see that stress effects measured by duration are consistent regardless of accent effect.[35] In Ortega-Llebaria (2006), the ratio values are smaller than 1.00, indicating that the vowel was longer than the consonant /g/ across the board. So, the C/V ratios in the stressed syllables in the present study decrease, rather than increase. This directional difference in the ratio results can simply be attributed to the varying degree of sensitivity to stress that is specific to segments. For instance, velar stops are intrinsically shorter than other stops, and also less susceptible to the suprasegmental effects in English (Umeda 1977). In the same study, Umeda showed that the absolute durational changes affected by prosody in duration are relatively greater for /s/ (40 ms ranges) than those of certain stops including velar stops (20 ms approximately). Thus, the durational changes in vowel were relatively greater than for onset [g] in Ortega-Llebaria's study while the durational changes of the fricatives were greater than those for the vowels in the current study. More important commonality between Ortega-Llebaria's study and the current result is that C/V ratio differences also can describe stress induced changes in syllables, and C/V ratio differences are found even at the level of dialects. Durational

---

[35] It should be noted that there are different conditioning factors involved between the current study and Ortega-Llebaria (2006): the current study differs from Ortega-Llebaria's in the number of syllables used in the target words (3 vs. 4), the placement of stress (initial or penultimate vs. final), and the consonant types (fricatives vs. stop). For example, we saw in Figure 3.4 that the syllables obtained C/V ratio values greater than 1.00 for most of the cases, indicating that the onset fricatives are longer than the vowels.

ratios between syllable constituents provide a useful way to examine stress effects on entire syllables and in a syntagmatic way within the syllable. That is, C/V ratio changes, rather than absolute durational changes, may be a way to signal relative prominence of stressed syllables compared to unstressed ones. As we show more results, we will be able to discuss on this issue.

We have seen in this section that durational differences are found between stressed and unstressed syllables (both onsets and vowels), indicating that duration is indeed one of the reliable stress correlates in Spanish. This durational result is in line with the view of stress correlates as independent phonetic manifestation (Sluijter & van Heuven 1996a; Sluijter et al. 1997; Ortega-Llebaria 2006; Ortega-Llebaria & Prieto 2011) when controlled for stress-accent covariation. That is, stress manifestation is distinct from other prosodic effects that can overlap in the temporal dimension, and variable duration can be observed for different segments. The examination of fricatives allowed us to find a consistent durational effect of stress on the consonants. This is in opposition to Campbell & Beckman (1997), who claim that stress itself may not have direct phonetic correlates. They investigated duration and spectral changes by stress in English, and found that durational changes are not significant for secondary stress. Their interpretation is that the apparent phonetic properties found in stressed vowels (duration and spectral tilt) are in fact affected by vowel reduction (rather than stress effects). The two different view regarding stress effects could be due to the language-specific ways of stress manifestation and at the same time different ways of prosodic interactions. As we show more results of other phonetic properties and the ways prosodic factors interact, we will be able to discuss further how we can better interpret the apparent differences.

The stress effect on duration described so far was based on a traditional way of measuring acoustic durations of consonant and vowel: marking the boundary between onset fricative and vowel at the voicing initiation in the waveform. This acoustic boundary may not be informative enough to understand what the temporal modulation would be within a segment. In order for us to better understand stress effects on syllable constituents, further acoustic/phonetic events were identified within consonant/vowel. As defined and discussed in Chapter 2, a more detailed type of measurement was also made (Narrow measurement) in identifying multiple

acoustic landmarks other than C and V boundaries. The following section examines the results of acoustic durations referring to the multiple acoustic landmarks.

3.2 Narrow measurement results

The narrow measurement specified in Chapter 2 uses multiple phonetic landmarks dividing the acoustic duration of consonants and vowels into further sub-segmental units. Integration of sequences of units, rather than each unit itself, will provide temporal details of the acoustic productions of the two fricatives [s, x] and the five phonemic vowels [a, e, i, o, u].

The segmentation included a total of four acoustic transitions before and after Consonant-Steady-State (C-ss) and Vowel -Steady-State (V-ss) portions. Thus, any syllable in this measurement consists of 6 contiguous sub-segmental intervals: C (t_1-C-ss- t_2) + V (t_3-V-ss- t_4). The summary of the phonetic landmarks used in the narrow measurement method was provided in Table 2.3 (Chapter 2). In the following sections, the investigation of subphonemic temporal behavior will enable us to understand how syllable durations can be characterized as a function of durational changes of C-ss, V-ss, and transitions. For this analysis, only the unaccented target words are considered to abstract the accent effect away from stress effects (N=1200).

3.2.1 Durations of consonant/vowel steady state (C-ss / V-ss)

In the narrow measurement, the duration of C consists of pre-consonantal transition (t_1), steady-state of C (C-ss), and post-consonantal transition (t_2). For the vowel duration (V), pre-vocalic transition (t3), steady-state of V (V-ss), and post-vocalic transition (t_4). The averaged durations of subcomponents of /s/ and V are summarized in Table 3.5.

Table 3.5 Averaged durations of the sub-phonemic units within the syllables (Narrow measurement)

Dialects	Stress	Duration (C = /s/) in ms.			Duration (V) in ms.			Duration (Syll) in ms.
		t_1	C-ss	t_2	t_3	V-ss	t_4	
Argentinian Sp.	Stressed	28.2(9.5)	76.7(23.7)	14.4(4.5)	13.7(5.8)	49.9(12)	20.2(9.5)	203.1(27.8)
	Unstressed	21.8(9.1)	65.7(20.2)	14.3(3.8)	13.2(6.6)	24.5(8.6)	18.4(7.8)	157.9(18.9)
Peruvian Sp.	Stressed	23.8(9.1)	87.3(26.6)	15.2(3.7)	20.4(7.9)	36.4(8.6)	23.3(8.6)	206.5(17.2)
	Unstressed	21.9(8.2)	68.1(19.0)	14.1(3.6)	14(5.5)	29(7.0)	15.5(7.5)	162.7(20.2)

The durations of C-ss and V-ss are longer for the stressed /s/ than for the unstressed ones. Stress as a main effect is significant for both C-ss ($F_{(1,9)}$=56.88, p<.001) and V-ss ($F_{(1,9)}$=264.02, p<.001) according to the Repeated Measures ANOVAs. This result is in line with the stress effect on C and V in the broad measurement method. The averaged durations of the /s/V syllables are shown in the last column of the table. As noted earlier, they are almost identical between the two dialects within each stress condition. The difference between the dialects is too small to reach a significance level ($F_{(1,9)}$=.69, p>0.05). However, when we compare C-ss and V-ss within stressed conditions, the two dialects differ in the stressed condition for both C-ss and V-ss in such a way that C-ss is longer in Peruvian Spanish while V-ss is longer in Argentinian Spanish. There are significant interactions between dialect and stress on the durations of C-ss ($F_{(1,9)}$=6.72, p<.029) and the V-ss ($F_{(1,9)}$= 54. 99, p<.001). Since the total syllable durations are similar between the two dialects, it is clear that stressed syllables are not simply produced by stretching the segmental components of the syllable proportionally. Instead, there are variable changes among the sub-segmental units including four transitions (labeled as $t_{1\sim4}$).

For representational purposes, the durations of /s/V syllables are visually summarized in Figure 3.6 in a compositional way.

Figure 3.6 Durational composition of the averaged /s/V syllables: stressed vs. unstressed
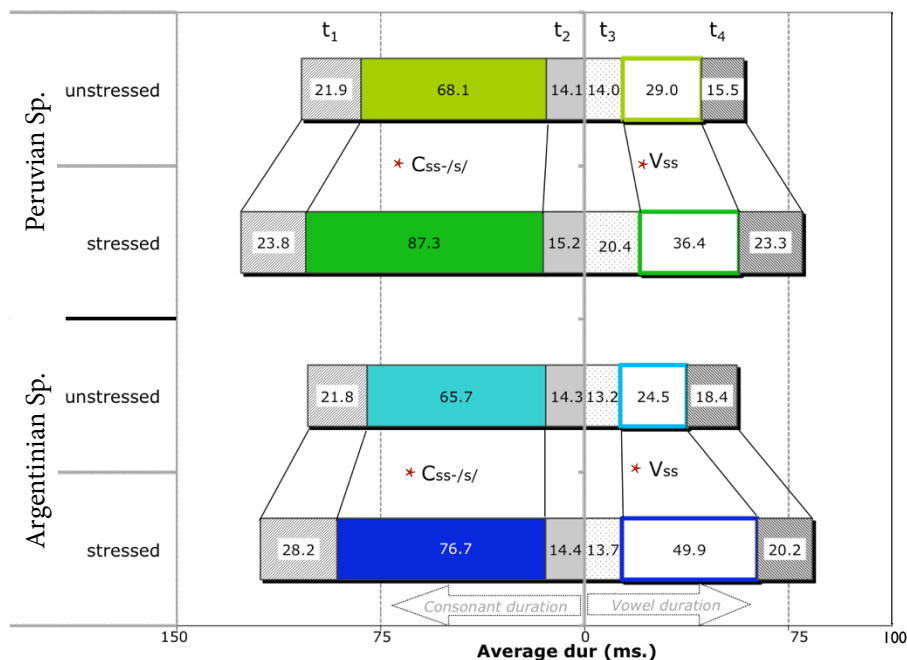
Note: The value "0" is aligned at the acoustic landmark, which is taken at the first peak of periodic cycles for a vowel. C-ss-/s/ refers to "Consonant Steady-State" in [s] and V-ss to "Vowel Steady-State" based on the measurement criteria detailed in Chapter 2.

The details of the durational composition reveal a systematic pattern that is characteristic to each dialect. That is, in the stressed syllables, the steady-state of the consonant (labeled as C-ss) is longer for Peruvian Spanish while the vowel steady-state is longer for Argentinian Spanish. In other words, in realizing stress, the durational change in the consonant steady-state is relatively stronger in Peruvian Spanish than in Argentinian Spanish, and vice versa for the vowel. Recall that durations of consonants and vowels taken in the broad measurement method were not significantly different as reported in the previous section. The narrow measurement results here report novel findings on the temporal characteristics between the two dialects. Overall, we see that the contribution of each syllable constituent (onset/rhyme) to stress manifestation varies between the two dialects; the consonant contribution is greater in Peruvian Spanish than in Argentinian Spanish, while vowel contribution is greater in Argentinian Spanish than in Peruvian Spanish. If we take into consideration that a durational change in V-ss plays a role in

cueing stressed syllables as prominent, the pattern that we find here becomes more interesting in that impressionistic description of Argentinian Spanish is supported by the results. That is, listeners would perceive Argentinian Spanish as being rhythmically more deviant than other dialects of Spanish due to the longer portion of the vowel steady-state. This could not be captured solely based on the broad measurement method of the segment durations.

A similar pattern is found in the /x/V syllables. The durational result is summarized in Table 3.6, and representative syllable durations are also displayed in Figure 3.7. The figure compares again the two dialects in the same way used in Figure 3.6 for /s/V syllables.

Table 3.6 The compositional duration of the averaged syllables by stress: /x/V syllables

Dialects	Stress	Duration (C =/x/) in ms.			Duration (V) in ms.			Duration (Syll) in ms.
		t_1	C-ss	t_2	t_3	V-ss	t_4	
AS	Stressed	**25.6**(10.6)	**65.2**(24.7)	**19.2**(6.5)	**15.1**(6.4)	**53.1**(15.3)	**20.8**(7.9)	**199.1**(30.4)
	Unstressed	**21.2**(9.1)	**55.8**(22.8)	**17.5**(4.7)	**14.2**(6.5)	**26.3**(9.2)	**19.8**(9.7)	**154.8**(29.2)
PS	Stressed	**25.93**(9.8)	**75.9**(25.7)	**16.5**(3.7)	**21.5**(6.5	**37.4**(11.4)	**22.1**(7.8)	**199.1**(31.8)
	Unstressed	**21.3**(8.3)	**58.7**(22.7)	**15.0**(3.9)	**17.7**(6.6)	**27.3**(6.8)	**16.9**(8.6)	**156.9**(25.0)

Figure 3.7 Durational compositions of the averaged /x/V syllables: stressed vs. unstressed
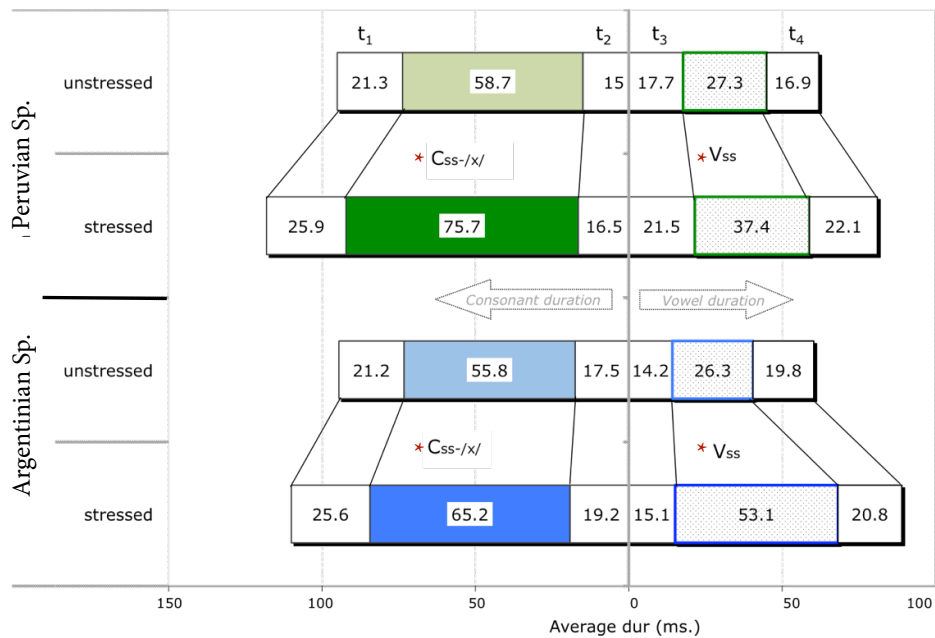
Within each stress condition, the averaged durations of /x/V syllables are almost identical between the two dialects: 199.06 ms (Argentinian Spanish=Peruvian Spanish) for stressed syllables and 154.81 ms (Argentinian Spanish) and 156.93 ms (Peruvian Spanish) for unstressed syllables. For the C-ss and V-ss durations, the asymmetric pattern is also found in /x/V syllables. That is, it is characteristic in /x/V syllables that C-ss in stressed syllables is longer in Peruvian Spanish than in Argentinian Spanish, and the durational pattern of V-ss is the opposite showing longer duration in Argentinian Spanish than in Peruvian Spanish. Therefore, the durational difference of C-ss-/x/ between stressed and unstressed syllables is greater in Peruvian Spanish (17 ms) than in Argentinian Spanish (9.4 ms) while that of V-ss is greater in Argentinian Spanish (26.8 ms) than in Peruvian Spanish (10.1 ms). This means that the consonant duration (C-ss) in Peruvian Spanish is influenced by stress to a greater degree compared to Argentinian Spanish, and the opposite pattern holds in the vowel durations.

Regarding consonant duration, we saw earlier that /s/ tends to be longer than /x/ ($p<.05$). As shown in Figure 3.6 and Figure 3.7, this durational difference comes mainly from a difference in the steady-states: C-ss-/s/ (87 ms) > C-ss-/x/ (76 ms) in Peruvian Spanish, C-ss-/s/ (77 ms) > C-ss-/x/ (65 ms) in Argentinian Spanish. For the unstressed condition, the difference is very weak but the same pattern holds. Repeated Measures ANOVA results confirm that the durational difference of C-ss by consonant type is significant for the stressed condition ($F_{(1,9)}=$ 19.74, $p<.002$), and not for the unstressed ($p=.09$). There was no significant interaction between the consonant types and stress on either C-ss or V-ss. Therefore, we can conclude that the durations of the two onset fricatives differ when stressed, but are similar when unstressed. As for the dialectal comparison, the presence of stress effects on C-ss/V-ss is consistently found, but the extent to which the durations of C-ss and V-ss vary by stress differs between the two dialects. This variable stress effect can be expressed via the ratios between C-ss and V-ss. Earlier in Figure 3.4, we saw how C/V ratios change in relation to the stress effect. We now show how the stress effects differ in the two dialects for the steady-state durations.

Figure 3.8 Stress effects translated into ratio
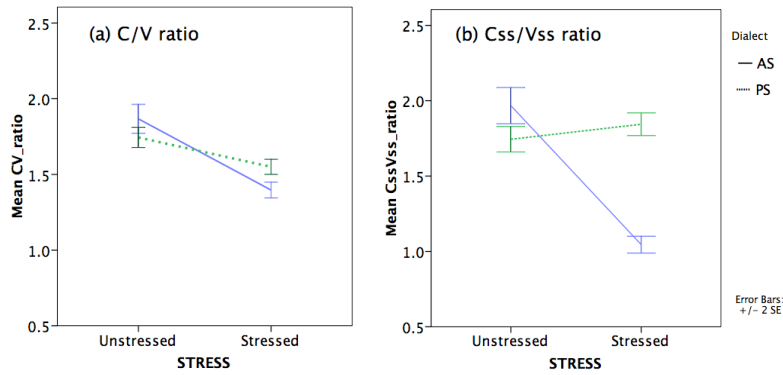
When we compare the C/V ratios between the dialects (on the left), we see that the difference exists weakly and only in the stressed syllable. This pattern becomes distinct between the two dialects when we compare the C-ss/V-ss ratios. When the syllable is stressed, the increase of the vowel duration, particularly V-ss, is much greater in Argentinian Spanish. This is different from the pattern in Peruvian Spanish, where the onset fricative increases as the vowel does, and even more increase in the onset is found to be characteristic of the stress effect for the Peruvian Spanish syllables. Thus, the two dialects can be described differently relative to the degree to which the C-ss and V-ss undergo the durational modification in the stress manifestation. This pattern is consistent between the two fricatives, as shown in Figure 3.9.

Figure 3.9 C-ss/V-ss ratio comparison between /s/ and /x/

65
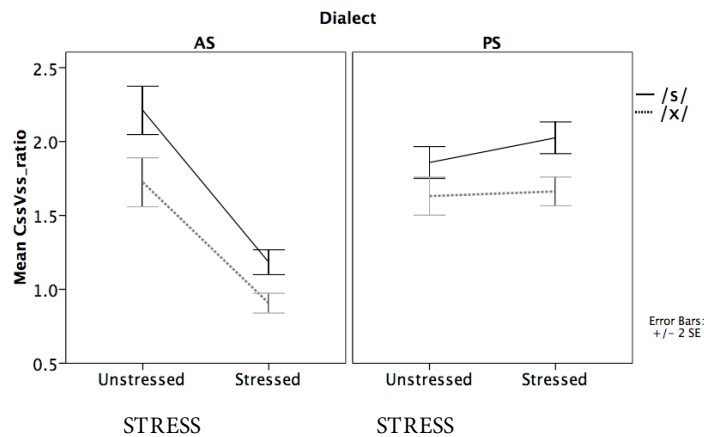
In this figure, we see decreasing C-ss/V-ss ratios on the left representing the stress effect in Argentinian Spanish. Relatively more stable ratios are characteristic of Peruvian Spanish on the right. This seems to suggest that the ratios between C-ss and V-ss effectively represent the stress effect for the given data. It also allows us to interpret how the syllable constituents integrate stress effects, and to compare variable stress effects possibly present in languages (or dialects). Our findings are in accord with Cho & Keating's (2007) articulatory results, which showed a stress effect in the (articulatory) seal duration of consonant but not in other measurements including acoustic duration. This means that stress effects can be reflected in sub-segmental units, which may not necessarily depend on the duration of segments (C and/or V). Thus, acoustic investigation of prosodic effects can be benefited from the narrow measurement method taken in the current, providing further details of phonetic information.

In the following section, we expand our investigation of stress effects to the behavior of transitions, and then we will be able to integrate the results of C-ss and V-ss in understanding the variable prosodic effects found in the dialects.

3.2.2 Transitions within the syllable

We now examine the durational patterns of transitions to understand the role they play in manifestation of stress and the dialectal variation that has emerged in the results discussed in the previous chapter. As we predicted, the total durations of transitions (t_1~t_4) are significantly different by stress ($F_{(1,9)}$=191.73, p<.001) and there is an interaction between dialect and stress ($F_{(1, 9)}$=16.34, p<.003). The question is whether the transitions also show systematic variation or whether it is simply because the total duration of the transition is a complement of the total syllable duration. Thus, we will examine how pre- and post- C-ss/V-ss transitions behave in the results. Assuming that the acoustic signal carries articulatory information to some extent, transitions may provide some information about movements, such as how fast the fricative constriction is formed and how quickly the vowel reaches its target. It is assumed here that a

shorter transition may signal a faster movement between two phonetic reference events when it is compared to a longer transition, all else being equal.

The mean differences of total transitions (T) between stressed and unstressed syllables are significant at the .05 level: $\text{Mean(T)}_{str}\text{-Mean(T)}_{unstr}$=15.1 ms. The narrow measurement method identifies four transitions within the syllable: pre-consonantal (t_1), post-consonantal (t_2), pre-vocalic (t_3), and post-vocalic transition (t_4) (refer to the sections 3.2.1 and 3.2.2 for the specifications of the transitions). In Figure 3.10, the transitions are compared by stress and then by dialect.

Figure 3.10 Durational summary of transitions in /s/V syllables
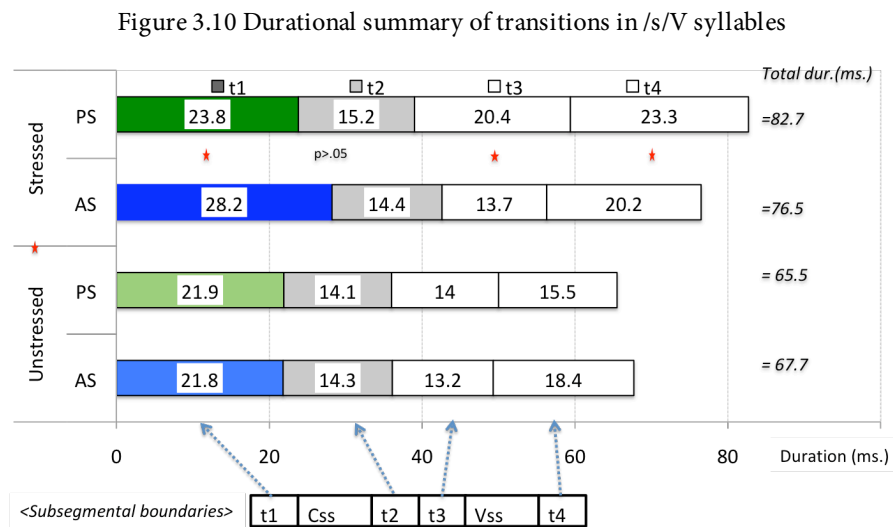
Notes: (a) The transitions are taken from /s/V syllables. (b) The asterisk in the figure represents statistically significant differences between the two dialects ($p<0.05$), and between the stress groups ($p<0.05$) according to the Repeated Measures ANOVAs.

The average durations of the transitions amounts to 66.5 ms and 79.6 ms for the unstressed and the stressed syllables, respectively, and the differences by stress are significant according to Repeated Measures ANOVA ($F_{(1,9)}$ =138.59, $p<.001$). In addition, there is a dialectal difference between the stressed syllables, but not between the unstressed ones. The total time spent on transitions is longer for Peruvian Spanish (82.7 ms) than for Argentinian Spanish (76.5 ms): STRESS×DIALECT, $F_{(1,9)}$=16.33, $p<.003$). Thus, the transitions are longer in stressed syllables

than unstressed syllables; for the stressed syllables, each transition, except t_2, shows a significant difference between the two dialects.

There are several interpretations that we can make from this transition result. First, the difference in the duration of C-ss between the two dialects is mainly in the different duration of t_1. Secondly, the interval between the release of the constriction and the onset of the voicing for the vowel is invariable (t_2), suggesting that glottal excitation for voicing is phased relative to the offset of C-ss. Thirdly, there are timing differences in reaching the vowel steady-state (t_3), and this makes the two dialects different in the narrow measurement method. Finally, the overall result of the transitions tell us about how stress manifestation incorporates the dynamic movements within the syllable, which is not simply to stretch each time domain to make syllables more prominent in signaling stress information. As mentioned earlier, the temporal prerequisite for transitions is influential in the perception of a consonant segment (Malmberg 1967).[36] This implies that listeners can tune in to such fine temporal modifications in interpreting acoustic cues. Thus, the transition analysis in this section demonstrates that such subtle variation can characterize languages even at the dialectal level, and that each dialect may have its own pattern of within-segment temporal variation.

Let us now summarize the durational changes depending on the changes in stress. As we have seen in the previous section, C-ss and V-ss showed significant differences by stress and also by dialect. The following figure shows how each sub-component of a CV syllable undergoes durational changes from unstressed to stressed syllables, and how such modifications differ between the two dialects.

---

[36] In Malmberg's (1967) study, the relative distance (or speed) of transition to vowel affected the listeners' perception of consonant (stop) affiliation either to the coda of the preceding syllable or onset of the syllable.

Figure 3.11 Durational dynamics within syllable
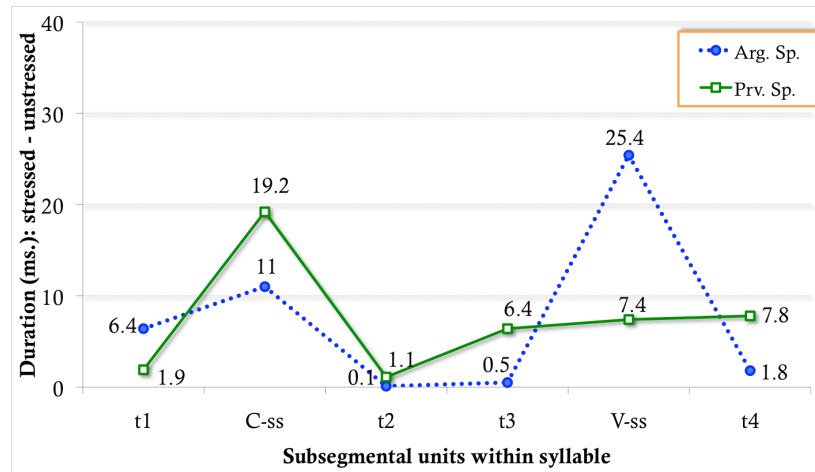
The most outstanding characteristic of this figure is that durational change within a syllable is asymmetrical between the two dialects. Peruvian Spanish, represented by the solid line, shows the greatest durational change within the consonant area (t_1, C-ss, and t_2) and fairly consistent durational changes in the transitions of vowel (t_3, V-ss, and t_4). Argentinian Spanish, represented with the dotted line, has most of the durational changes in the vowel. In particular, for Argentinian Spanish, the stress effect is greatest in the vowel steady-state (V-ss). This asymmetrical effect of stress on duration shows that there are variable ways to make a stressed syllable prominent in the temporal dimension. The C-skewed pattern found in Peruvian Spanish could be interpreted to indicate that the formation of a constriction for the fricative is reached relatively earlier/faster and stays longer compared to the other pattern, the V-skewed pattern found in Argentinian Spanish. On the other hand, the V-skewed pattern reaches the vowel target faster and stays longer, as reflected in a longer steady-state of vowel (V-ss).

The following figure shows the proportional contributions of transitions to the total durational changes affected by stress. There are four representative components in the figure within the syllable: transitions within C (t_1~t_2), transitions within V(t_3~t_4), C-ss-/s/, and V-ss.

69

Figure 3.12 Proportional durational changes by stress: unstressed unit as a reference
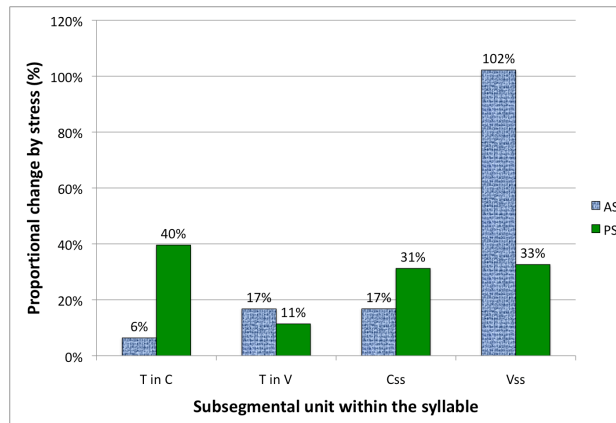
If the stress effect were to stretch an unstressed syllable simply proportionally over the syllable as a whole, we would expect the proportional changes of each part would be quite similar over the sub-segmental units; for example, the durational changes from the unstressed to the stressed syllables might be expected to be 25% more or less for each part (T in C, T in V, C-ss, and V-ss). However, this was not the case, and the durational changes were concentrated mostly on the steady-state of C and V. The total durational changes by stress were 44.6 ms for Argentinian and 46.3 ms for Peruvian. The proportional changes (%) in the figure were calculated for the stressed syllables with respect to the unstressed ones. For example, V-ss duration in Argentinian Spanish changed by 26 ms when stressed, and the proportional change is 102% since the unstressed V-ss is 25.4 ms. Each unit is calculated in the same way, and displayed in Figure 3.12. If the total duration by stress is simply a proportional stretch, we would expect to see similar degree of changes for each unit, whose durations are variable in the unstressed condition. For Peruvian Spanish, it is close to what we would expect, showing fairly similar amount of changes in each unit. However, Argentinian Spanish shows a dominant proportional change in V-ss, and small changes in the other units. This means that the stress effect is mostly concentrated on V-ss in Argentinian Spanish. This pattern clearly rejects the idea that stressed syllables are merely a stretch of unstressed syllables. Furthermore, temporal modulation clearly can be specified differently between dialects.

Recall that the two dialects showed an asymmetric pattern regarding which one of the segment parts (C-ss or V-ss) becomes longer in manifesting stress. Argentinian showed more increase in the vowel and Peruvian in the consonant. The fact that the two dialects showed similar durations of the total C and V suggests that the inter-segmental (C-V) modulation was similar between the two dialects, but the intra-gestural modulation (within C and V) was different. This dialectal variation raises an important question about how prosodic effects can properly be implemented in a linguistic/gestural representation through which variable speech output can be predicted in languages and dialects. We will explore a possible gestural account for the variation in Chapter 5.

To summarize, the detailed measurement of syllable duration provided us two interesting patterns that are consistent across the syllable sequences. First, the acoustic steady-states of the consonants and the vowels are the parts that undergo greater durational changes driven by stress. This shows that stressed syllables are not realized just by a temporal stretch relative to the syllable form, but by undergoing a *systematic* temporal modulation. Second, acoustic transitions are also involved in the durational pattern of stress implementation in such a way that a longer acoustic transition correlates with a shorter duration of the vowel steady-state. This pattern was consistent for /s/V and /x/V syllables. Further discussion will continue in the summary section 3.3, after examining how other factors may interact with stress effects, and also in Chapter 5 discussing a possible account for this variation.

Earlier in this section, we reported that there was a significant correlation between C duration and V-type. This may mean that the durational pattern we have found so far could be a local characteristic of certain vowel contexts rather than a fully general effect. In addition, we also found that different segments may behave differently in terms of stress effects, and this could also be the case for the vowel contexts. Thus, the investigation of the effect of vowel context will ensure that the variable stress effects are rather general across the vowel contexts.

3.2.3 Stress effects and vowel contexts

The speech data we have analyzed so far include the five phonemic vowels (/i, u, e, o, a/) in Spanish, with the results pooled across vowel contexts. However, vowels are known to have intrinsic durations, which correlates with the openness of the vowel: the lower the height, the longer in duration (Peterson & Lehiste 1960). Thus, the vowel duration could be expected to follow the order of /i, u/ < /e, o/ < /a/, all else being equal. Of interest here is whether the varying degree of prosodic effects that we have found so far is consistent regardless of the vowel identity.

Literature has shown that it is possible for different segments to be manifested in duration to varying degree. For example, in Fry (1955) we find different duration ratio results for the vowels in English noun-verb pairs that are contrastive in the stress placement (e.g., n. *pérmit* and v. *permít*): e.g., 'permit (noun:verb)' showed the ratios of 120:60 ms for the /ɚ/ whereas 120:90 ms for the /i/.[37] We will first describe the durational differences conditioned by vowel height, and how stress effects behave differently or similarly among the vowel height categories, and then check the consistency of the asymmetric pattern between the two dialects.

In Figure 3.13, the durations of C-ss and V-ss are compared for each vowel height category. Repeated Measures ANOVA results are also provided in Table 3.7.

Table 3.7 Repeated Measures ANOVA results: Vowel height effects on the C/V/Syllable durations

Tests of Within-Subjects Contrasts							
Source	Measure	Type III Sum of Squares	df	Error (VH)	Mean Sq.	F	Sig.
Vowel Height (VH)	V_total	868.415	1	10	868.415	58.431	0.000*
	V-ss	174.132	1	10	174.132	30.302	0.000*
	C_total	80.929	1	10	80.929	5.631	0.039*
	C-ss	63.867	1	10	63.867	3.019	0.113
	Syllable	419.138	1	10	419.138	11.575	0.007*

---

[37] The perception results indicated that the overall listeners' judgments corresponded linearly and less variably to the intensity ratio variation than their responses to the durational variation, which showed a sharp increase between 2 and 3 in the ratios. This implies that the perceptual sensitivity to the physical cues may function differently: that is, duration and intensity are involved in the stress manifestation, but differently: duration is a more categorical effect (at least for English).

Figure 3.13 Mean durations of C (C-ss) and V (V-ss) in /s/V syllables: pooled across stress conditions
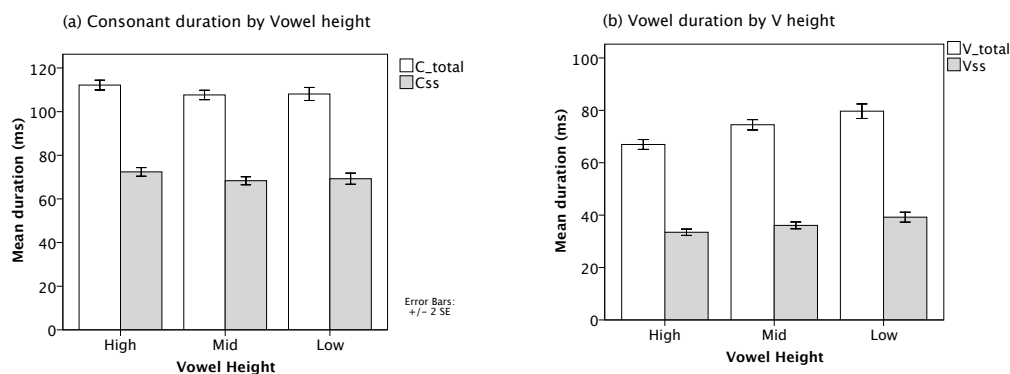
The vowel durations in (b) show a clear distinction among the vowel height categories; the lower the vowel height, the longer in duration. The V-ss durations also follow the vowel height hierarchy: that is, the duration of V-ss of a lower vowel is longer than that of a higher vowel. The durational differences intrinsic to the vowel height hierarchy are preserved regardless of the stress condition: that is, the vowel height has a main effect in the vowel durations even in the unstressed syllables ($p<0.001$). The consonant durations in (a), on the other hand, show relatively stable durational behavior. There is a vowel height effect on the C duration such that the fricatives preceding higher vowels show relatively longer consonant duration. Although this indicates a compensatory relationship in the durations between onset fricative and the vowel within the syllable, the vowel height still has a main effect on the syllable duration ($p<0.007$). When we separate the syllable durations by the stress conditions, we see that the vowel height effect is present only distinguishing the low vowel from the others in the stressed syllables. This is shown in Figure 3.14.

Figure 3.14 Syllable duration as a function of vowel height
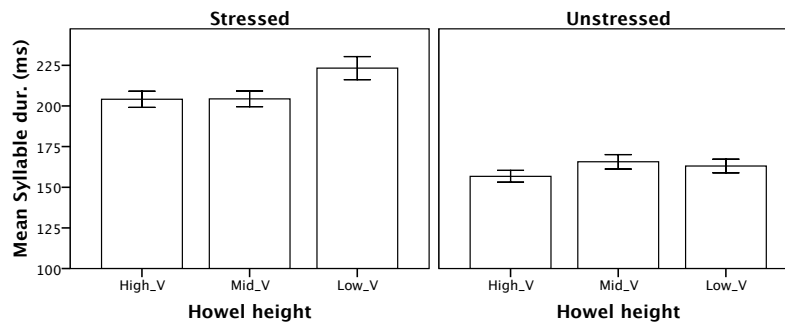
Both Scheffe and Bonferroni Post-hoc tests indicate that the syllable durations are different only between the low and the high vowel categories: the mid vowels are not statistically different either from the high or the low vowels.[38] This means that the syllable durations are more distinct among the vowel height categories (high & mid vowels vs. low vowel) when stressed.

Now we are in a position to examine whether the asymmetric effect of the stress between the two dialects is consistent among the vowel height groups. Recall that the stress effect was strongest on the C-ss for Peruvian Spanish and on the V-ss for Argentinian Spanish. The mean differences in duration between the stressed syllables (V-ss-str) and the unstressed (V-ss-unstr) syllables are plotted over the sub-segmental units in Figure 3.15.

Figure 3.15 Stress effects by vowel height

(a) Argentinian Spanish by vowel height   (b) Peruvian Spanish by vowel height

---
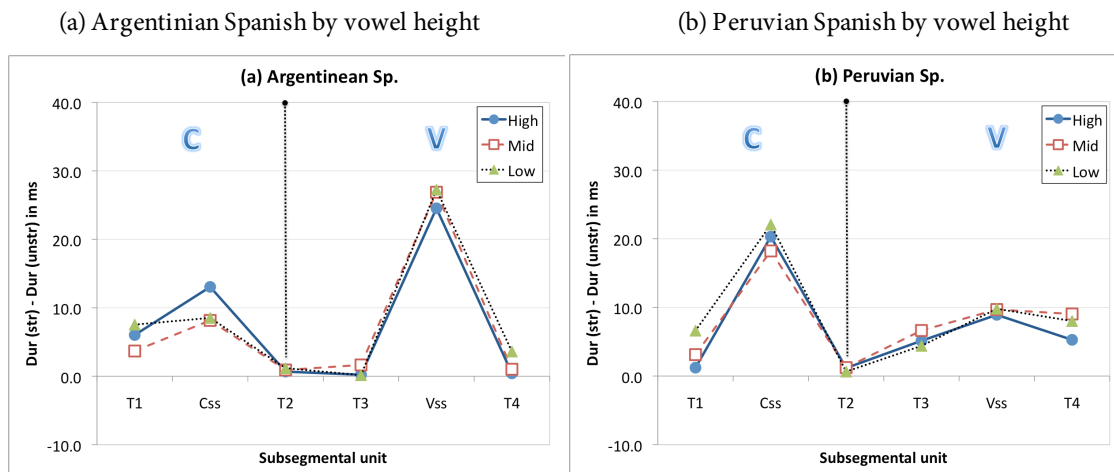
[38] Equal variances are not assumed (Dunnett's C) since the low vowel group (/a/) have less tokens than the other two groups (/i,u/ and /o, e/).

Comparing figures (a) and (b), we see that the asymmetric pattern is clearly present within each vowel group (high-mid-low). Stress effects are the greatest in the V-ss for the Argentinian stressed syllables in (a), and the three vowel categories follow the same pattern. Similarly, for the Peruvian stressed syllables in (b), the consonant lengthening is the most prominent change by the stress condition, and again the stress effect on the C-ss occurs to a similar extent among the vowel categories. Thus, the vowel context varied in the data is not involved in the pattern variation between the two dialects, and we can safely conclude that the asymmetric durational pattern is characteristic of variable stress effects on the syllable constituents, and that the pattern is indeed the variation found in the stress manifestation between the two dialects.

3.2.4 Position in word effects

Prosodic boundaries often constrain the realization of segmental events, and levels of prosodic boundary to which segments are sensitive, and/or degree of sensitivity to boundary effects seems to vary among languages (Delattre 1966; Hoequist 1983a; Cho et al. 2002; Keating et al. 2003; Cho 2005). Our main interest here lies in determining whether this effect, particularly Position in Word (PoW), interacts with stress, and if so, whether the effect is consistent across dialects.

PoW is known to influence the phonetic and/or phonological realization of segments, particularly for onset consonants. Although stress and PoW are strong factors for different reasons, the manifestation of both effects can overlap in the acoustic dimensions.[39] For example, Umeda (1977) examined English consonant durations, and concluded that PoW and lexical stress were the most influential conditions for the durational variation that she found. The duration of /s/ was longer (129 ms) word-initially and the same (120 ms) word-medially, and the effect was also found in the unstressed /s/. On the other hand, Lavoie (2001; pp.112) reported that the durations of /s/ were identical between word-initial and medial positions, for both

---

[39] For example, initial position in word is claimed to be psychologically strong and prominent since it has an important role in lexical access and word recognition, compared to later parts of the word (Cole & Jakimik 1980; Cutler & Norris 1988; Marslen-Wilson & Zwitserlood 1989; Beckman 1998)

stressed /s/ (120 ms word-initially and medially) and unstressed /s/ (104 ms word-initially and 107 ms word-medially), even though she found the PoW effect for other segments (e.g., stop consonants).[40] Although the durational results between the two studies differ regarding the PoW effect for the fricative, they agree that stress and PoW effects can condition variable durations of segments. We examine now whether the PoW effect is present in the fricatives /s/ and /x/, and whether the PoW effect interacts with stress in Spanish.[41]

The target words were all extracted from IP medial position. This was to minimize potential overlapping with other sources of prominence on the target words, such as focus and pitch accent (Gabriel 2006). The CV syllables in the target words differ in PoW. For example, the target syllable /-so-/ is word-initial in the word _sótano_ and word-medial in _pasóta_. In both words, the target syllable /so/ is stressed, and differs minimally in the PoW condition. Similarly, unstressed syllables also have counterparts that differ only in PoW. An examination of the PoW effect will ensure that the stress effect we find is independent of the PoW effect.

Repeated Measures ANOVA results are summarized in Table 3.8; we will discuss PoW related effects marked in bold in the table.

---

[40] Lavoie (2001; p.95) discussed the variation pattern of sibilant, and noted that sibilant fricatives showed the least variation across positions in her results. She discussed a possible reason for this pattern referring to the observations made in the literature: "… sibilants require more articulatory control" (Ladefoged & Maddieson 1996; 137).

[41] Lavoie (2001; pp. 95) described that the stress effect on /s/ is weaker relatively to the other consonants she investigated: "In both the initial and medial position, we observe very little variation relative to stress (for the sibilants investigated)."

Table 3.8 Within-subjects effects of vowel and syllable duration.

Dependent Variable	Independent variable (factors)		F	Sig.
I. C-duration	a.	**PoW**	$F_{(1,9)}= $ **4.047**	**p = 0.075**
	b.	C_type	$F_{(1,9)}= 7.12$	p < .05*
	c.	Stress	$F_{(1,9)}=77.039$	p < .001**
	d.	**PoW × Stress**	$F_{(1,9)}=$**4.881**	**p = 0.055**
	e.	**PoW × C_type**	$F_{(1,9)}=$**6.304**	**p < 0.05***
	f.	PoW × Stress × C_type	$F_{(1,9)}=2.4$	p = 0.156
	g.	**PoW × C_type × Stress × Dialect**	$F_{(1,9)}=$**4.775**	**p = 0.057**
II. V-duration	a.	PoW	$F_{(1,9)}= 0.007$	p = 0.937
	b.	*C_type*	$F_{(1,9)}= 4.24$	*p = 0.07*
	c.	Stress	$F_{(1,9)}= 222.999$	p < .001**
	d.	**PoW × Stress**	$F_{(1,9)}= $**56.141**	**p < .001***
	e.	PoW × C_type	$F_{(1,9)}= 1.309$	p = 0.282
	f.	PoW × Stress × C_type	$F_{(1,9)}= 0.72$	p = 0.418
	g.	PoW × C_type × Stress × Dialect	$F_{(1,9)}= 0.046$	p = 0.835
III. Syll-duration	a.	PoW	$F_{(1,9)}= 2.529$	p = 0.146
	b.	C_type	$F_{(1,9)}= 2.069$	p = 0.184
	c.	Stress	$F_{(1,9)}= 211.638$	p < .001**
	d.	PoW × Stress	$F_{(1,9)}= 0.828$	p = 0.386
	e.	**PoW × C_type**	$F_{(1,9)}= $**4.623**	**p = 0.06**
	f.	**PoW × Stress × C_type**	$F_{(1,9)}= $**3.641**	**p = 0.089**
	g.	**PoW × C_type × Stress × Dialect**	$F_{(1,9)}= $**4.651**	**p = 0.059**

We have seen in the previous section that there is a stress effect on the durations of C and V, and also on the duration of the syllable as a whole (I-c, II-c, and III-c). Besides the stress effect as a main effect, we find interactions of the PoW with Stress (I-d), C-types (I-e), and Dialect (I-g, III-g), each of which will be discussed in the following.

First, Figure 3.16 shows how the mean durations of consonants, vowels, and syllables differ with changes in PoW.

Figure 3.16 PoW effects on segment durations across the dialects
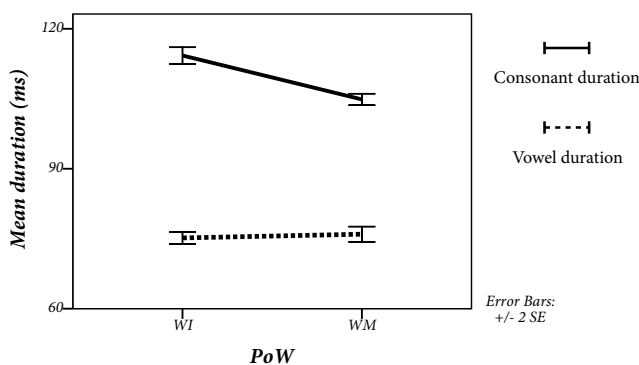
The mean durations appearing in Figure 3.16 are the averaged durations pooled across stress and accent conditions. We see that the onset is sensitive to PoW but not for the vowel. In general, it has been found that initial consonants tend to be longer and greater in articulatory magnitude, and temporal modulation is also expected to occur, if PoW effects are present (Byrd 1994; Keating et al. 1999).[42] For example, Kuzia et al. (2007) found that word-initial fricatives in German are acoustically longer in duration and becomes more 'fortis-like' with the increased glottal vibration. Similar effects are found in various languages; e.g., word-initially longer VOT in Korean voiceless stops (Jun 1993), and strengthening of /h/ (Pierrehumbert & Talkin 1992).[43] In line with these studies, our duration result can be interpreted to mean that the PoW effect is one of the conditioning factors in the durational variation of the fricatives, albeit a weak effect.

An examination of how stress interacts with the PoW effect reveals that the longer durations are found word-initially for the onsets (*p*=.055) but word-medially for the vowels (*p*<.001). The following figures show how stress and PoW interact each other.

Figure 3.17 Interaction between stress and PoW: duration comparison

(a) C duration by PoW   (b) V duration by PoW   (c) Syllable duration by PoW

---
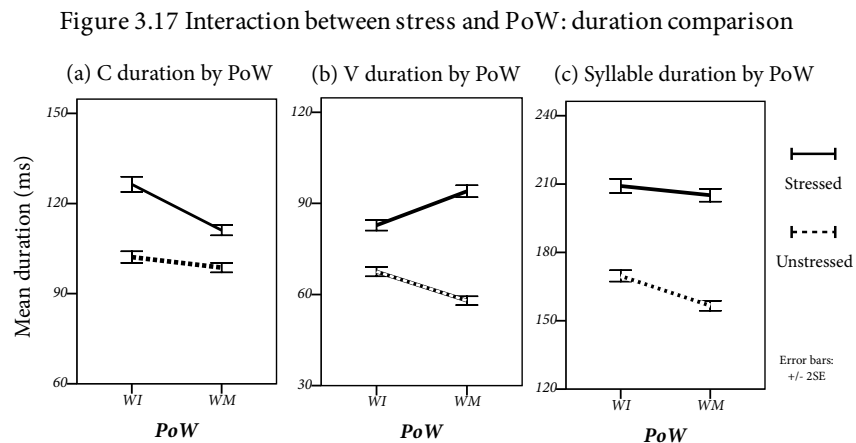
[42] However, Cho & Keating (2007) found that initial vowels are also affected by stress and increase intensity. We will discuss this result when we show our intensity result.

[43] Interestingly, Tabain et al. (2001) also found an interaction between acoustic segment durations and the prosodic hierarchy in French, although their results showed that consonant durations, particularly for fricatives, were rather stable/unaffected by the boundary effect, compared to those of vowels. This result is not directly comparable to ours since the consonants and the vowels are located in different syllables (e.g., no syntagmatic interpretation within the syllable is possible).

The mean duration is longer word initially than medially for the stressed consonant in (a), but this initial lengthening remains as a tendency for the unstressed consonant. Interestingly, the opposite pattern is found for the syllable duration in (c); that is, the duration is not different by PoW for the stressed syllables but is different for the unstressed syllables. This means that the consequence of PoW effects on each syllable constituent results in similar durations of stressed syllables regardless of the word position. It shows that the stressed vowel is longer in (b) when the stressed onset is shorter word-medially in (a). On the other hand, the stressed vowel is shorter in (b) when the stressed onset is longer word-initially in (a). This durational relation between the onset and the vowel suggests a possible complimentarity within the syllable in relation to word position, and this relation holds only in the stressed syllable. Thus, our interpretation is that the complementary results of C and V durations are induced by stress since this relation does not appear in the absence of stress.

The asymmetric pattern of the PoW effect reminds us of the asymmetric pattern between the two dialects regarding the durational behavior of C-ss and V-ss: the duration of C-ss was longer in Peruvian Spanish and that of V-ss was longer in Argentinian Spanish for the stressed fricatives. The presence of the PoW effect raises a question of whether the longer durations of C-ss and V-ss are mainly attributable to the stress effect, or the pattern results from variable PoW effects. If the dialectal difference lies in the stress effect independently from the PoW effect, the asymmetric pattern in the durations of C-ss/V-ss are predicted to persist regardless of position. Figure 3.18 confirms that this is the case.

Figure 3.18 Comparison of the PoW effect by stress: C-ss and V-ss
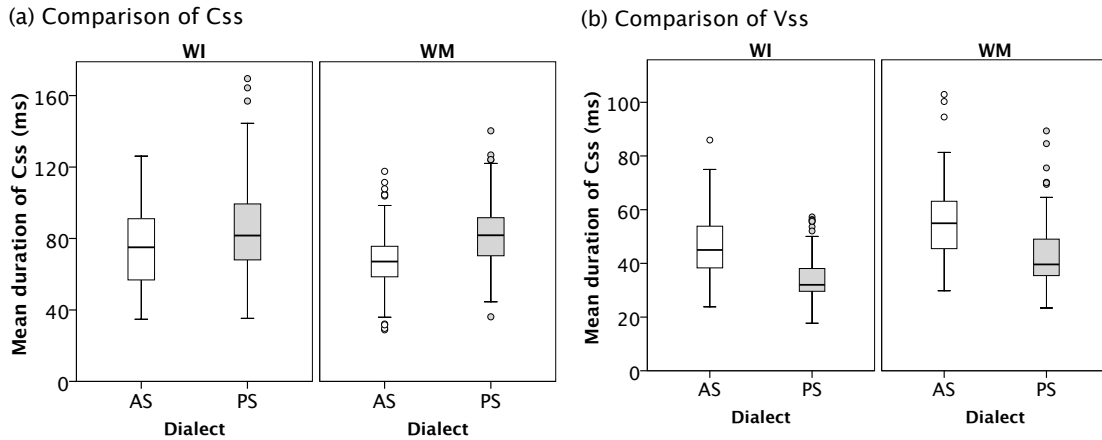
The comparison made in the figure considers the stressed syllables only since the unstressed ones were similar between the two dialects in both stress and PoW effects. The durational difference of C-ss is retained between the two dialects in both the positions (a), and the difference increases word-medially. This indicates that the PoW is involved in the word-initial position, and that the durational difference between the dialects is modified by the PoW effect. On the other hand, the durational difference of V-ss is similar regardless of the positions, and the durations increase word-medially to a similar extent for the two dialects. The different patterns between the C-ss and V-ss are in accord with the facts that the PoW effect is to lengthen the duration of the onset, and that the vowel shortening seems to be compensatory for the onset lengthening within the syllable.[44]

The dynamic changes in sub-segmental duration are compared between the initial and the medial positions in Figure 3.19. We see that the characteristic durational differences are present regardless of the PoW. That is, the durational modification in the C-ss is greater in Peruvian Spanish than in Argentinian Spanish, and vice versa for the V-ss. The figure is based on the stressed syllables only.

---

[44] One way to understand this compensatory shortening of the vowel is that languages do show rhythmic compensation as reading styles become regularized.

Figure 3.19 Durational dynamics and the PoW effect: stressed syllables only
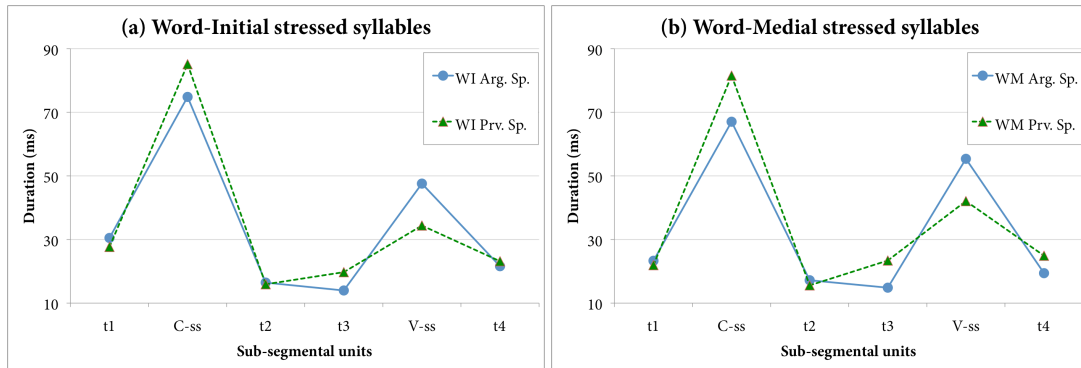
One-way ANOVA was conducted to compare the two dialects for the C-ss and V-ss durations. The results show that the dialectal difference in the C-ss is marginally significant ($p=.059$) and significant for the V-ss ($p<.05$). In addition, the transition from the voicing onset to the vowel steady-state is shorter word-initially in both dialects, and the difference in T3 between the two dialects turns out to be significant according to One-way ANOVA ($F_{(1,599)}=159.48$, $p<.001$). This indicates that the vowel steady-state in the stressed syllables is achieved earlier and retained longer in the Argentinian Spanish than in the Peruvian Spanish. This is not the case for the consonant, where the transitions before and after the consonant steady-state seem to be stable and the only difference between the two dialects is how long the steady-state is maintained for the consonant articulation.

The PoW effect is also compared between the two fricatives to see whether the place of articulation is involved in the PoW effect. In the earlier discussion of PoW effects, we noted that previous studies suggest that different segments may or may not show PoW effects. For example, PoW effects were present in stops but not in the fricative for Tabain et al. (2001). When fricatives do show PoW effects, coronal fricatives seem to be least sensitive to the effect (Lavoie 2001). For example, Lavoie (2001; 115) included the durational differences in Spanish fricatives by the PoW, and the results show that the durations of the coronal fricative /s/ differ from the fricatives in other places. The duration of /s/ was identical to 101 ms (WI=WM) for the pre-stress condition, while it was 101 ms (WI) and 84 ms (WM) for the non-pre-stress condition. On the

other hand, /β/ (76: 67), /f/ (108: 101), and /x/ (120:106) all showed positional differences in duration: these durational differences by PoW were consistent in the non-pre stress condition.

Contrary to the previous study, when we compare the durations of the fricatives in the current data, we find that the PoW effect is present in /s/ but not in /x/. Figure 3.20 shows that the PoW effect is present in /s/, but not in /x/ based on the result of the broad measurement method.

Figure 3.20 PoW effects for the two fricatives: /s/ vs. /x/
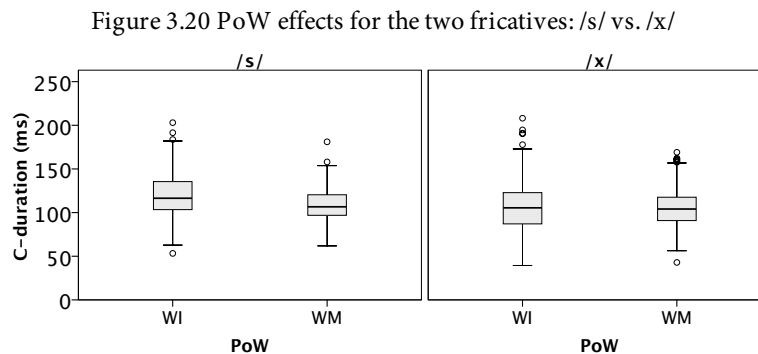
Repeated Measures ANOVA results indicate that the PoW effect on /s/ is significant (p=.023), but not for /x/ (p>.05). A second interpretation is borne out when we examine the interaction between stress and PoW effects over the sub-segmental units based on the narrow measurement. That is, the PoW effect is clearly present except the unstressed /x/ in Argentinian Spanish. Thus, this inconsistent durational pattern of /x/ between the two dialects is involved in the overall result of the PoW effect between the two fricatives. This is shown in Figure 3.21, where the results are further separated by the dialects.

Figure 3.21 PoW effects comparison between /s/ and /x/: durational changes in sub-phonemic units
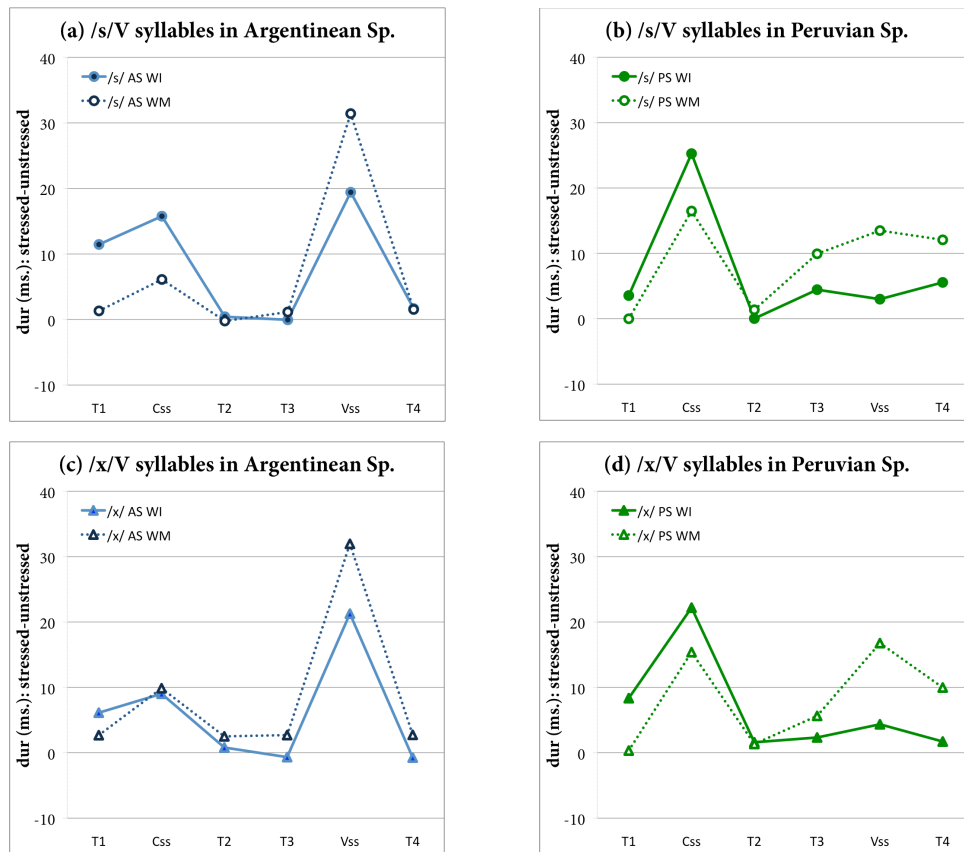
Note that the durational differences between stressed and unstressed C-ss-/x/ are relatively smaller for Argentinian Spanish (a and c) than for Peruvian Spanish (b and d). This seems to suggest that the durations of C-ss-/x/ are less sensitive to both stress and PoW effects. The overall mean durations of /x/ are indeed shorter in Argentinian Spanish (M=110.05 ms) than those in Peruvian Spanish (M=119.88 ms) when the fricative is from stressed syllables. When /x/ is unstressed, no difference is found (94.5 ms for Argentinian Spanish /x/ and 95.2 ms for Peruvian Spanish /x/). Except for the behavior of the Argentinian Spanish-/x/, all the other cases show PoW effects such that the word-initial C-ss is longer than the medial counterpart. Thus, the apparently different durations from the multiple factor interactions (PoW × C-type × Stress ×

Dialect; I-g and III-g in Table 3.8) are attributable to the dialectal difference for the durations of /x/.[45]

To summarize, the PoW effect interacts with stress and is present in the durations of both onset fricatives and vowels: stressed C duration is longer word-initially while V duration is longer word-medially. This asymmetrical behavior of C and V influences the total duration of syllables in a complementary fashion, and results in no difference in the syllable duration as a whole.

3.3 Summary of stress effects in duration

We have seen so far how the durations of C and V are influenced by stress, and how the stress effect interacts with other factors such as accent, PoW, onset C-type, and V context in the phonetic realization of stress. Our questions were how stress affects the durations of C and V within and across the two dialects, and how the stress effect interacts with, or is influenced by, other linguistic factors such as vowel types (by height) and PoW. In the following sections, we will briefly review the results combined from the two measurements (Broad and Narrow), and then discuss some of the advantages born out from the narrow measurement method, focusing on variable prosodic effects manifested in different dialects. In particular, dialectal variation will be discussed in terms of temporal modulation of syllable and variability.

3.3.1 Duration result summary

We found that stressed syllables are significantly longer than unstressed syllables in the two dialects of Spanish studied here, Argentinian and Peruvian Spanish. In addition, both onset and vowel were significantly longer when stressed, indicating that both constituents are indeed active in the temporal modulation of stressed syllables.

---

[45] Alternatively, it could be due to the difference in the phonetic stability of /x/ between the two dialects.

The durational patterns of onset fricatives can be summarized as follows: 1) there was a segment intrinsic difference in duration between the two fricatives (/s/>/x/). 2) Accent effects are found to be selective: both stressed and unstressed /s/ are longer when accented, but not for /x/. 3) In the unaccented condition (IP-medially), consonant duration changes by stress were greater for /s/ than /x/. This is in line with the previous result in which stress effects on the durations of stop closure varied for different place of articulation (Umeda 1977). This result was further investigated between the two fricatives, and it was concluded that the apparently weak stress effect on /x/ is due to the overall short duration of /x/ in Argentinian Spanish when stressed. In other words, the stressed /x/ in Argentinian Spanish is not longer in duration compared to its unstressed counterpart, and as a result the overall stress effect on /x/ is seen as inconsistent and weaker. 3) The durations of onset consonants were also influenced by PoW such that stressed fricatives became longer word-initially than word-medially. Unstressed consonants, on the other hand, were not significantly sensitive to the PoW effect, although the effect still remains as a tendency. This result indicates that there is a significant interaction between stress and PoW effects in the data investigated in this study.

Vowel durations were found to be more sensitive to stress effects. Firstly, the overall vowel durations were different by vowel height probably due to the intrinsic vowel variation. The intrinsic durations held for both the unstressed and the stressed vowels. This means that durational differences of unstressed vowels are not neutralized by the stress effect. Secondly, stress effects on the vowel durations correlated with vowel height, and this was particularly the case of Argentinian Spanish, which showed more active stress effects on the vowels: /i, u/ < /e, o/ < /a/. This means that the stress effect can be greater as intrinsic vowel duration is longer.

Consonant durations showed an inverse correlation with vowel height: in particular, fricatives followed by high vowels were longer in duration compared to those in the non-high vowel contexts. One way to understand this pattern is that there can be a rhythmic compensation between the syllable constituents in the temporal dimension. In other words, stressed syllables can retain similar durations by the durational compensation of the onset. This rhythmic compensation is analogous to the situation where number of syllables in the word tends to influence syllable durations; that is, syllable duration tends to decrease as the number of syllables

in the word increases. As Harris & Umeda (1974) discussed, word compression effects can occur when utterances are obtained in a regularized speaking mode such as using a carrier phrase.[46] It can be speculated that the apparent durational difference of the fricatives before the high vowels occurs because the onset fricatives are lengthened in order to compensate for the intrinsically short vowel duration. This has the effect of accommodating rhythmic regularity at a lower level, here, the syllable.

The durational changes of syllable constituents were translated into C/V ratios. The result showed that C/V ratio with the fricative onset (/s/ or /x/) decreases from unstressed to stressed contexts. This was because the proportional increase in duration was greater in vowel than in consonant. Between the two dialects, the changes of C-ss/V-ss ratios were greater in Argentinian Spanish than in Peruvian Spanish from the Narrow measurement results. This means that similar ratios between C-ss and V-ss tend to be preserved regardless of stress condition in Peruvian Spanish, while a drastic increase/decrease of the ratios occurs in Argentinian Spanish, due to the greater durational changes in vowel as a function of stress when compared to Peruvian Spanish. This was interpreted to mean that the vowel duration is more affected by stress in Argentinian Spanish compared to Peruvian Spanish, where a more balanced effect is found between the constituents in manifesting stress.

The intra segmental variation found in the temporal modulation of stress is a novel finding that invokes an interesting question of how these fine-detailed modulation patterns could be encoded in linguistic systems. In addition, a question arises regarding what the underlying

---

[46] Harris & Umeda (1974) investigated the durational differences of vowels depending on speaking mode, and found that the vowel durations do not vary in a connected speech mode: in other words, no word compression effect is found. They explore a reason why different speech modes have an impact on temporal modulation of segments, and suggest that the role of rhythm differs in different speech modes. In a more regularized utterance, such as using a carrier phrase, rhythmic tempo is more dominant, and compression effect can emerge. On the other hand, speakers in a connected speech mode use more dynamic prosody that conveys the message.

articulatory mechanism would be in order to account for this variation. A gestural account in order to accommodate this variation will be provided in Chapter 5.

3.3.2 Methodological summary

This section summarizes the duration results from a methodological point of view. The acoustic segment durations were measured in two ways: one referred to voicing initiation to mark the boundary between onset consonant and vowel, and the other subdivided acoustic events into transitions and steady-states of each segment. The second method supplements the first one by having multiple acoustic landmarks to detail temporal events that might be involved in stress manifestation, such as movements (and speed of movements) between steady-states. The purpose of this section is to integrate the results borne out from the two methods employed in the measurements, and also to highlight some of the advantages taken from the narrow method.

First, in section 3.1, we saw that segment durations are not significantly different between the two dialects based on the broad measurement, except that consonant duration was marginally longer in Peruvian Spanish than in Argentinian Spanish. When we referred to the Narrow measurement results, we could see that vowels as well as consonants were different between the two dialects in terms of steady-state portions. This result could support the impressionistic description of Argentinian Spanish having an unusual rhythmic pattern compared to other dialects of Spanish.

Secondly, one of the findings from the Narrow measurement method is that the stress effect on duration is clearly present in the steady-state portions of both C and V in a consistent way. In addition, the way of implementing stress influences the temporal structure of syllable even at the level of dialects. The durational change driven by stress was more concentrated on the onset in Peruvian Spanish, and on the vowel in Argentinian Spanish. This could mean that syllable constituents may contribute variably to the syllable prominence driven by stress, leaving the question of what motivates this variation and how this pattern can be accounted for.

The third finding is that the C-ss/V-ss ratio is more representative of dialectal variation than the full C/V ratio. The new ratio result could describe the characteristic difference between the two dialects such that C-ss and V-ss ratio is more stable, regardless of stress, in Peruvian Spanish than in Argentinian Spanish. The ratios increased more drastically as a function of stress in Argentinian Spanish, suggesting that there is a more durational change involved in Argentinian Spanish regarding the stress effect.

Finally, it was found that the stress effect does not just involve stretching the time-span relative to an unstressed syllable. The degree to which onset or vowel duration changes by stress varies between the two dialects such that C-ss increases in duration more in Peruvian Spanish than in Argentinian Spanish, and vice versa with the duration of V-ss. This asymmetric pattern, interestingly enough, results in similar durations of the syllable on the average. One interpretation of this is that there could be a quasi-complementary relationship between the syllable constituents in the manifestation of stressed syllables. In other words, the contributions of syllable constituents to stress prominence could vary by language (or dialect). In addition, combining the two results, one way we can understand the asymmetric pattern is to say that the temporal modification of C and V could presumably refer to a higher-level prosodic constraint, such as isochronous rhythm that is characteristic of Spanish. In other words, syllable constituents are actively involved in making stressed syllables more prominent in the temporal dimension but they also have to preserve the upper limit regulated by the rhythmic characteristic. This inference can be made since the two dialects were similar in duration of syllables within each stressed condition. Related discussions and possible accounts will be explored in Chapter 5 after we examine spectral and intensity characteristics of syllables when stress is involved.

More insightful results were discussed based on the Narrow measurement employed in this study. The narrow method enabled us to better understand the temporal modification influenced by stress (within the acoustic dimension), and also to capture micro-variation at the level of dialects. These advantages suggest a way to overcome the limit of the traditional approach of measuring acoustic durations, for which temporal investigations of acoustic signals have often been undermined in modern decades. As many studies have pointed out in literature, phonetic properties of a sound segment are not always clear by marking C and V only, and sub-contrastive

or gradient phonetic phenomena in speech cannot be captured in such a way. In addition, considering the fact that some perceptual cues are present in the transitions between two segments, it is possible that speech production may utilize such details resulting in systematic variation across languages/dialects. For example, Stevens (2007) documented preaspirated stops in Sienese Italian, and found a role of preaspiration in the perception of gemination by native speakers. That is, native speakers of Sienese Italian make a linguistic interpretation from the existence of preaspiration (e.g., preaspiration as part of consonants). Similarly detailed phonetic investigations into dialects can help us establish phonetic/phonological connections of sub-segmental variation to a linguistic system. More ways to develop dynamic temporal measurements deserve more attention to acquire higher consistency in performing technical measurements.

Chapter 4 Intensity and spectral modulation by stress

This chapter examines qualitative aspects of stress effects by assessing changes in intensity and spectral distributions found in the onset fricatives and the vowels. The intensity measurements in section 4.1 include overall intensity levels as well as RMS (Root Mean Square) ratios for the fricatives and the vowels as a function of stress. In the following two sections, two spectral measurements will characterize stress effects on the energy distribution of the fricatives and the vowels, respectively. The spectral characteristics of prosodic effects on the two fricatives are examined in section 4.2, by measuring four spectral moments (Center of gravity, variance, skewness, and kurtosis) in the onset fricatives. Energy distributions over frequencies of vowels are examined in section 4.3 using measure of spectral balance above 0.5 kHz using four filtered bands (B1~B4). The investigation of spectral distributions in the fricatives and the vowels enables us to define stress effects beyond static characteristics that refer to a mean value within a phonetic parameter. Finally, vowel formant changes as a function of stress are also investigated in section 4.4. Section 4.5 summarizes the intensity and spectral results.

According to Sluijter & van Heuven's (1996a,b) studies, intensity as one of the stress correlates can be better understood by examining the energy changes across frequencies, namely spectral balance, rather than overall intensity levels. They found that only the intensity at higher frequencies changed as a function of stress, and not the lower energy (below 0.5 kHz). Spectral balance as a stress correlate has been confirmed by other studies: American and British English (Turk & Sawush 1997; Turk & White 1999), Catalan (Astruc & Prieto 2006), and Spanish (Ortega-Llebaria & Prieto 2006). Stress effects found in spectral balance can be understood as having greater vocal efforts that affect higher energy. For instance, Glave & Rietveld (1975) examined effects of varying effort on the spectral intensity distribution, and found that greater intensity was distributed at higher frequency region (above 0.5 kHz) of vowels when the speakers

put greater vocal effort in their speech production. Under the assumption that stress realization correlates with vocal effort, we would expect to see the stress effect in our data.

Two main questions will be addressed in this chapter. Firstly, how is stress manifestation within the domain of the syllable reflected in the intensity and spectral dimensions? Secondly, how is the stress effect in each dimension distinct from other prosodic effects such as accent and PoW effects, and how do the prosodic factors interact with one another. The disentanglement of other prosodic effects from the stress effect will clarify what is the fundamental phonetic modification introduced by stress, and how prosodic factors interact and are interrelated in the realization of speech in general. The investigation in this chapter will lead us to a more complete picture of stress effects by incorporating the temporal characteristics investigated in the previous chapter.

4.1 Intensity as a function of stress

We use RMS (Root Mean Square) ratio between syllable constituents as a way to compare stress effects within the syllable as a function of stress and other prosodic factors.[47] We find a RMS C/V ratio analysis in Ortega-Llebaria's (2006) study, where she described consonant strength relative to the vowel duration and found that C/V ratio in Spanish increased in the unstressed syllables regardless of the presence/absence of pitch accent. The C/V ratio was also known to play a role in determining voicing of postvocalic consonants (Denes 1955; Port & Dalby 1982). That is, the vowel duration relative to the consonant duration is involved in the voicing identity of the consonant. This suggests that a relative measure can be useful in examining the syntagmatic relation of segments. Thus, the C/V ratio measure as an indication of relative consonant strength

---

[47] RMS is directly related to the power in the signal. A speech signal transformed into waveform can be viewed as time-varying energy distribution. Amplitude describes this energy on the waveform in the first transform. However, the absolute values in amplitude are not linearly correlated to our perceptual energy of speech signal, as many studies have reported in the phonetics literature. For example, Stevens (1987) emphasized the need of acoustic measurements as relative terms rather than absolute measurements.

can describe the relation between syllable constituents that may vary depending on prosodic conditions.

The RMS values were obtained over the entire onset consonant and vowel, respectively. C-RMS is divided by V-RMS and the ratio is mostly less than 1 because vowel amplitude tends to be much greater than that of consonant in general. Accordingly, an increase in RMS ratio can mean either that consonant energy increases proportionally more than the increase of vowel energy, or that energy decreases more for the vowel than for the consonant. In this calculation, a greater RMS ratio is positively correlated with relatively stronger consonant or weaker vowel. In this respect, the C/V ratio comparison between stressed and unstressed syllables will serve as an indicator of a relative consonant strength with respect to that of vowel. Since C/V ratios in the RMS results depend on two variables (C-RMS and V-RMS), we first examine RMS values in the onset fricatives and the vowels in section 4.1.1, and then we discuss the ratio values in section 4.1.2.

4.1.1 Stress effects correlate mainly with vowel RMS

This section aims to determine whether stress effects are reflected in the energy level of the onset fricatives and/or the vowels. In general, sibilant fricatives are expected to have greater overall energy compared to non-sibilant, and the overall RMS result confirms that this is also the case in our data: the mean RMS value is greater for /s/ (0.15 Pa) than /x/ (0.10 Pa) [$p<0.001$, according to One-way ANOVA analysis]. In addition, Within-Subjects analyses show that Accent and Stress as main factors did not have a main effect on C-RMS, but did have a main effect on V-RMS [Accent: $F_{(1,9)}=20.08$, $p=0.002$; Stress: $F_{(1,9)}=25.85$, $p=0.001$]. There was also a significant interaction between Accent and Stress on both C-RMS and V-RMS: C-RMS [ACCENT×STRESS: $F_{(1,9)}=10.466$, $p=0.01$] and V-RMS [ACCENT×STRESS: $F_{(1,9)}=13.19$, $p=0.005$]. That is, the RMS values of both the C and V increase as a function of stress in the presence of accent, while only the V-RMS increases in the unaccented condition. This interaction is shown in Figure 4.1.

Figure 4.1 Stress effects on RMS: interaction with Accent

(a) C–RMS

(b) V–RMS
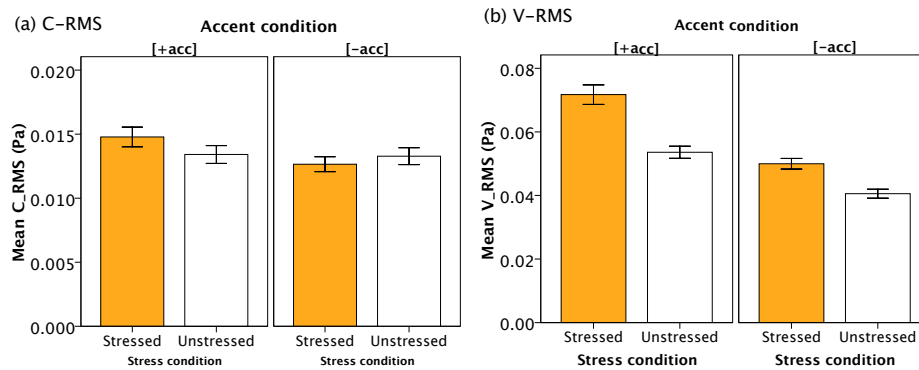
In the accented condition, the means of the consonant RMS differ between stressed and unstressed fricatives [$F_{(1,1199)}$=6.98, p=0.008]. The means of vowel RMS differ between stressed and unstressed conditions regardless of accent. This indicates that the stress effect may be magnified in the onsets of the accented syllables. In addition, the stress effect on the vowel RMS is present regardless of the presence of absence of accent, although the effect is reduced in the unaccented position. There was no significant interaction between the onset fricative type and stress /accent.

Since the stress/accent effect in the intensity dimension is found to be asymmetric between C and V, and the overall intensity of syllables depends mainly on the stress/accent effect on the vowel. The overall stress/accent effect over the domain of syllable, as shown in Figure 4.2 below, can be understood such that the magnitude of the prosodic effect depends mainly on how much stress/accent influences the V-RMS (Figure 4.1b).

Figure 4.2 Stress and accent effects on the syllable RMS

93
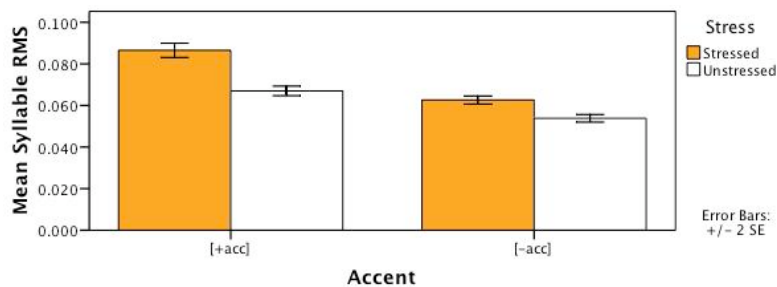
Stress and accent as independent factor have main effects in the syllable-RMS according to the within-subject results of Repeated Measures ANOVA: STRESS [$F_{(1,9)}$=17.19, p=0.003] and ACCENT [$F_{(1,9)}$=25.96, p=0.001].

The asymmetric stress effect between C-RMS and V-RMS may be connected to the temporal results. We have seen earlier in Chapter 3 that the durational changes influenced by stress were not equally proportional between onset and vowel. The vowel lengthening was proportionally greater than the consonant lengthening. This may suggest that temporal increase correlates with RMS such that longer duration results in greater RMS. If we consider vowel durations among the vowels by height, we can predict that low vowels would be greater in the RMS results. We examine the durational increase by vowel height, and compare it to V-RMS increase. The results are shown in Figure 4.3.

Figure 4.3 Correlation between duration and intensity of vowel
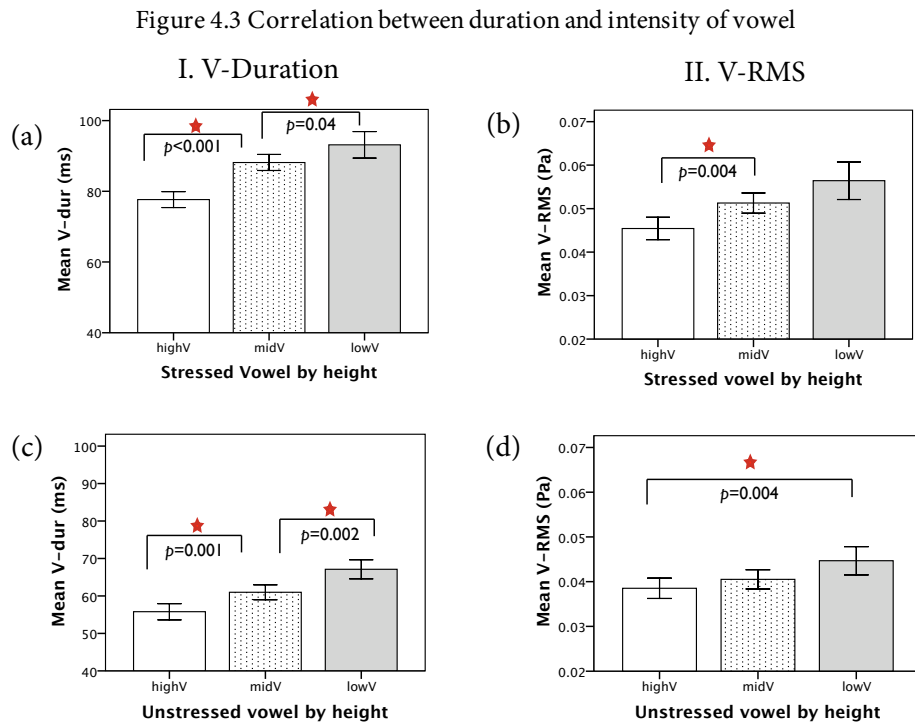
In these two figures, the vowel duration and RMS increase as vowel height changes from high to low. When the vowels are unstressed (bottom figures), the durational contrasts among the vowel height categories are still retained as shown in (c). On the other hand, the V-RMS differs only

between highV and lowV, which suggests that duration and intensity are not necessarily correlated intrinsically. The statistical results shown in the figure are from Bonferri post-hoc tests (95% significance level). When the vowels are stressed (top figures), the durational contrast is retained in (a), and the V-RMS becomes additionally contrastive between highV and mid-V. The additional contrast in the V-RMS can be understood as the stress effect, which is greater as the vowel height lowers. We can interpret this result in two aspects. First, duration and intensity is not tightly linked, though tendency can be found. Second, the stress effect in the intensity dimension can be greater as vowel duration increases. In particular, the mid vowel lengthening by stress is the greatest in (a), and the intensity of the mid vowel is also the greatest as in (b). Thus, it can be interpreted to mean that intensity and duration are not intrinsically correlated, but the stress effect can be greater as the vowel lengthens more.

Now we turn to PoW effects, which may be seen as a boundary effect. One may expect to see greater energy word-initially, since the PoW effect is considered as one of the domain-initial strengthening (Keating et al. 2003). However, the result indicates that this may not be the case in the intensity dimension. The overall PoW effect in RMS is found in the consonant RMS: [PoW on C-RMS: $F_{(1,9)}$=16.48, p=0.003; PoW on V-RMS: $F_{(1,9)}$=0.09, p=0.77]. In contradiction to our prediction, the RMS is greater word-medially than word-initially, as shown in Figure 4.4.

Figure 4.4 Consonant RMS by PoW
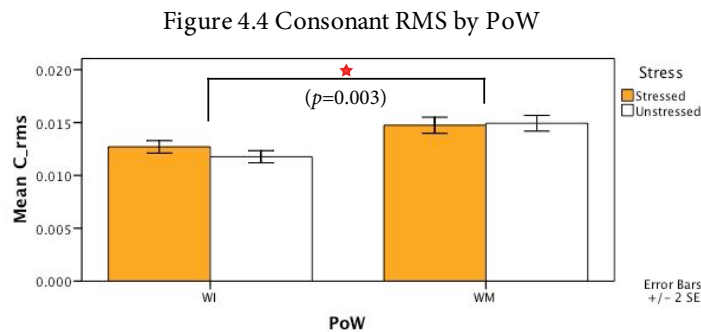
This result provides two important implications. First, the PoW effect in the intensity dimension is the opposite pattern of the stress/accent effect that we have seen earlier. That is, the PoW effect is mostly reflected in the consonant intensity while the stress/accent effect was mainly on the vowel intensity. This suggests that PoW is a different kind of prosodic effect from the

stress/accent effect, particularly regarding the prosodic domain even within the syllable. The second implication is that the overall intensity in RMS may be a poor indicator of the positional effect, compared to other parameters such as duration and spectral changes. In the intensity dimension, the PoW effect is not a strengthening word-initially. Thus, our interpretation is that the PoW effect needs to be understood as a modification in the temporal and spectral dimensions, and not in the overall intensity dimension.

An additional implication is borne out when we examine the interaction between PoW and accent conditions: [$F(1,9)=25.45$, $p=0.001$]. The position effect in the consonant RMS is sensitive to the presence and absence of accent such that the effect is only present in the accented condition, as shown in Figure 4.5.

Figure 4.5 The PoW effect on C-RMS
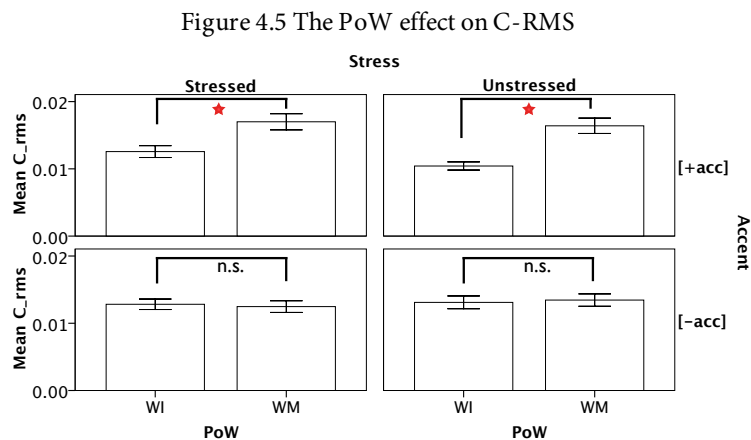
When the fricatives are accented, they show greater RMS values word-medially regardless of the stress condition. PoW as an independent factor has a main effect: PoW[$F_{(1,9)}=16.48$, $p=0.003$]. No difference of the PoW effect is found between stressed and unstressed conditions. The PoW effect conditioned by the presence of accent is found consistent between /s/ and /x/, though the degree of the effect vary between the two fricatives (/s/ > /x/). The result implies that the PoW effect may not even be involved in the intensity modification. Instead, the positional intensity difference reflects the sensitivity of accent effects to the PoW.

4.1.2 RMS C/V ratio as relative strength of onset

Since overall energy difference could also be influenced by non-linguistic conditions, such as loudness of individual production and recording conditions, we examine the results using the C/V ratios in the syllables. The C/V ratio result will allow us to interpret how relative consonant strength within the syllable is influenced by prosodic factors. The mean C/V ratios pooled across conditions are 0.25 (SD=0.15) for the stressed syllables and 0.31 (SD=0.17) for the unstressed syllables. The smaller ration in the stressed syllables is interpreted to mean that the vowel RMS is greater relative to the consonant RMS. The mean C/V ratios are smaller in the accented condition than those in the unaccented. Thus, decrease in C/V ratio of RMS is a characteristic of both stress and accent effects. Finally the PoW effect when translated into C/V ratio shows the same pattern with the stress and accent, as shown in Figure 4.6.

Figure 4.6 Prosodic effects translated into C/V ratios
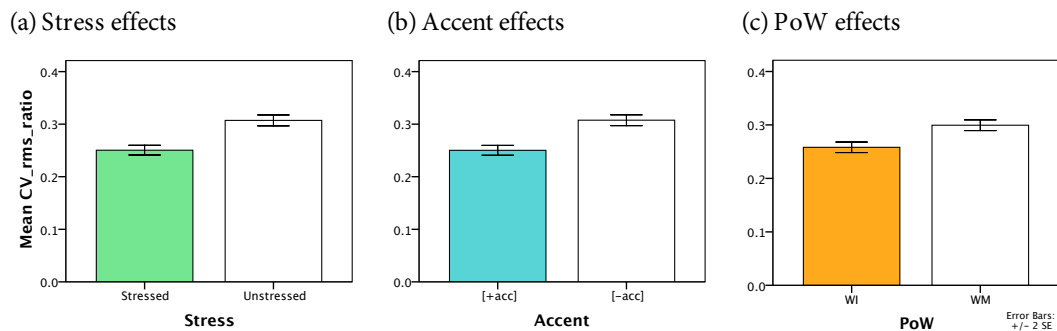
The C/V ratio translation presents a consistent pattern among the three types of prosodic effects such that the presence of each effect is associated with a smaller ratio. This can be interpreted to mean that the consequence of the prosodic effect is the relatively greater intensity of the vowel either by increasing vowel intensity (i.e., stress/accent effects) or by reducing the consonant intensity (i.e., PoW effects).

This pattern is consistent between the two fricatives, as shown in Figure 4.7.

Figure 4.7 C/V ratio: stress and accent effects
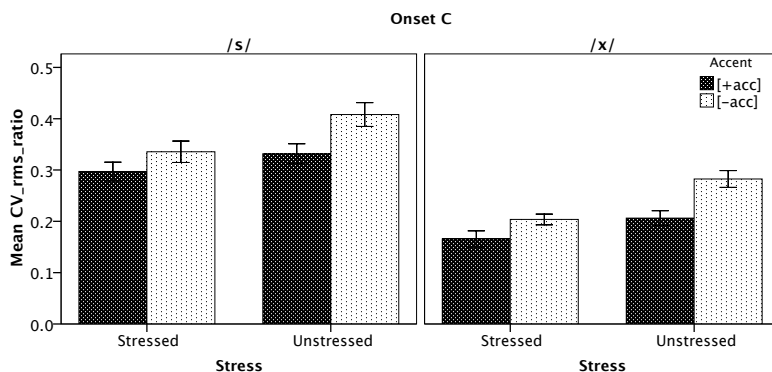
In these figures, the C/V ratios decrease when the syllables are stressed and/or accented, and the pattern is consistent between /s/V and /x/V syllables. The only difference is in the magnitude of the ratios such that the /x/V syllables show smaller ratios compared to the /s/V syllables. The decreasing C/V ratio pattern among the three prosodic factors is also consistent within each onset condition. As we see in Figure 4.8, the only difference is these ratio comparison is that the /x/V syllables have smaller ratios in comparison to those for the /s/V syllables across the three types of prosodic effects (a~c).

Figure 4.8 Prosodic effects translated into C/V ratio: consistency between /s/V and /x/V syllables
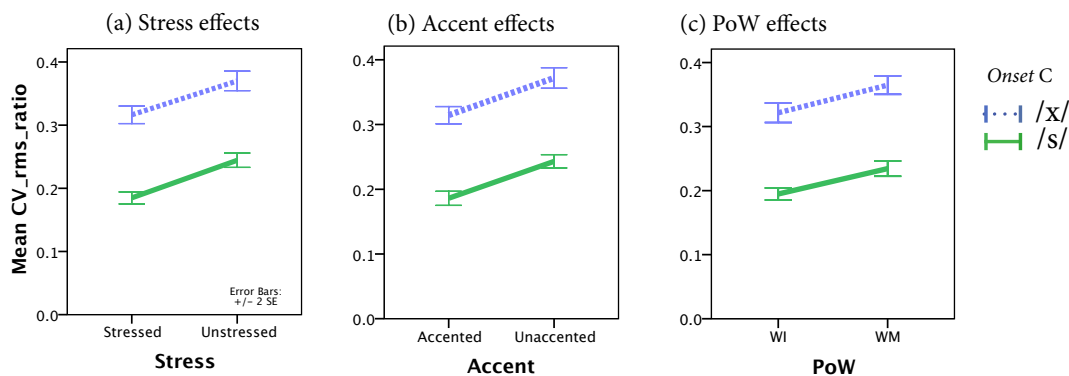
These figures confirm that the C/V ratios are stable in describing each prosodic effect, even between the /s/V and /x/V syllables. Syllables influenced by each prosodic effect consistently show lower C/V ratios, and the ratios for the /x/V syllables are lower than those of the /s/V syllables. This suggests that the C/V ratio analysis in intensity can be a reliable description of the

prosodic effect in the intensity dimension regardless of the variable onset compositionality.[48] A final note on the PoW effect is that this word-boundary effect is constrained by accent conditions. That is, when syllables are unaccented, the PoW effect becomes absent. The following figure shows this conditional effect of PoW.

Figure 4.9 PoW effects in C/V RMS ratio conditioned by accent: pooled across onset contexts
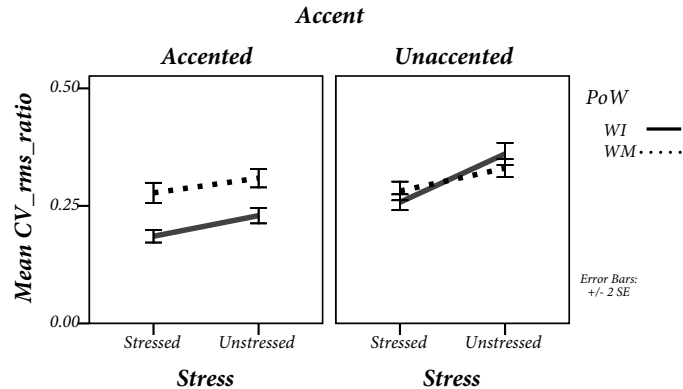
In this figure, word-initial syllables show smaller C/V RMS ratios regardless of stress conditions when they are accented (left panel), but this PoW effect disappears when the syllables are unaccented (right panel). This suggests that word position can be constrained by other prosodic factor (e.g., accent in this case). This will be discussed further in Chapter 5 when we compare individual and combinatory prosodic effects.

A dialect comparison is also made to see how the C/V ratio stability can help us understand the intensity behavior between the two dialects. The overall RMS values were higher in Peruvian than in Argentinian Spanish for consonant, vowel, and the syllable: DIALECT on C-RMS [$F_{(1, 2,399)}$=58.85, $p<0.001$], on V-RMS [$F_{(1, 2,399)}$=61.11, $p<0.001$], and on Syllable-RMS [$F_{(1, 2,399)}$=77.99, $p<0.001$]. This RMS difference does not have any linguistic interpretation, and it can be trivial. However, the C/V ratios between the two dialects are identical, as shown in Figure 4.10.

---

[48] The C/V ratio pattern is also consistent between the two dialects.
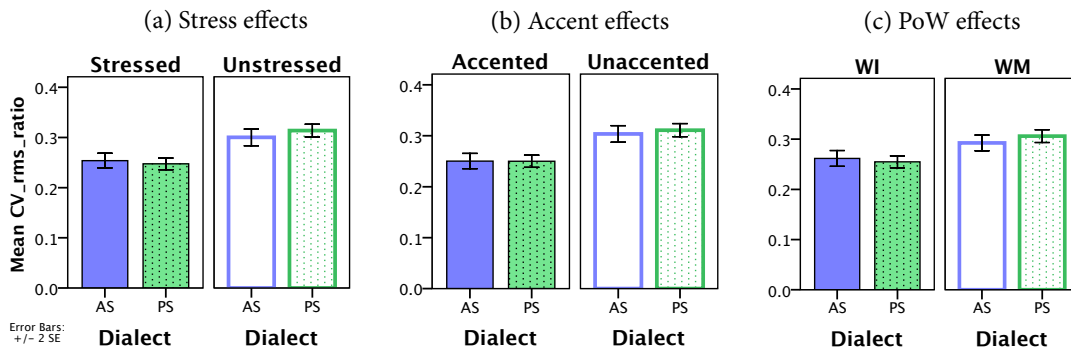
In these figures, the dialectal differences that we saw earlier in the absolute RMS values of C, V, and syllable, disappeared across and within conditions. This C/V ratio analysis now describes that the two dialects are homogeneous in the intensity behavior.

To summarize, this section investigated stress effects in the intensity dimension as measured by the RMS energy. Stress influenced vowel energy to a greater degree than the onset fricatives. The accent effect was also similar to the stress effect such that accented syllables show greater energy than the unaccented syllables. We also examined PoW effects and it turned out that the effect is mainly involved in the consonant energy, which is different from the other two prosodic effects. However, the C/V ratio analysis revealed that all three prosodic effects resulted in the same pattern: that is, a smaller C/V ratio is associated with the syllables that are more prominent (e.g., stressed, accented, and word-initial syllables). In addition, the apparent differences found between the two dialects in the intensity dimension become interpretable such that the prosodic manifestation in the intensity dimension is indeed homogeneous in the two dialects.

4.2 Stress effects in the onset fricatives as a measure of spectral moments

This section is concerned with stress effects in onset fricatives as a measure of four spectral moments. The stress effect on the onset fricatives will be evaluated by comparing the changes in mean frequency (also known as Center of gravity; M1), standard deviation (M2), skewness (M3),

and kurtosis (M4) of the fricative spectra. The four moments are expected to detail spectral characteristics as a function of stress by means of the spectral energy distribution referring to the mean frequency and the values of skewness and kurtosis. Since the spectral characteristics of onset and vowel are distinct, no direct comparison will be made between syllable constituents. However, the spectral behaviors of each constituent will shed light on the connection between temporal and spectral behaviors in the manifestation of prosody.

For the Center of gravity measures of the fricative spectra, a higher COG is expected for fricatives with a shorter anterior resonating cavity: thus, COG for /x/ should be considerably lower than the COG for /s/. Spectral outcome of fricative production depends partly on the size and shape of the oral cavity in the vicinity of a fricative constriction, through which turbulence noise is generated and the source noise at the glottis is filtered (Fant 1960; Stevens 1998). Spectral Center of gravity is used in describing different places of articulation for fricatives (Jassem 1979; Forrest et al. 1988; Nittrouer et al. 1989; Ladefoged & Maddieson 1996; Shadle & Mair 1996; Behrens & Blumstein 1998; Jongman et al. 2000; Gordon et al. 2002). However, recent studies have shown that stress effects, or prosodic effects in general, can be reflected in the spectral Center of gravity. For example, Cho & Keating (2009) found a prosodic effect as well as measured by COG: stress influences COG of stop burst and, though weaker, accent seems to affect COG of consonants.

The four spectral moment analyses will help us identify spectral changes induced by stress in the onset fricatives. One of the main predictions is that onset fricatives will show higher COG when they are stressed, accented, and/or located word-initially. We will examine stress effects in the fricatives as measures by COG, as well as the three other spectral moments (variance, skewness, and kurtosis). The advantages of the spectral moment analysis are as follows: the second moment, variance, can describe how overall energy distribution is spread from COG at the frequency domain. The third and the fourth moments can be used to describe whether more energy is tailed toward left (lower frequencies) or right (higher frequencies) from the mean frequency, and how heavy the tails of the distribution would be over the frequency domain.

In order to identify the stress effect proper, we will abstract out potential effects from the other conditioning factors, such as prosodic positions (within the IP and Word). In the following two sections, we examine accent and PoW effects, and then turn to stress effects.

4.2.1 Accent effects in the spectral moments depend on the onset types

Accent effects conditioned by the IP position were first investigated by means of four spectral moment comparisons between accented and unaccented positions. The averaged COG values (first moment: M1) were 6.32 kHz IP-initially and 5.94 kHz IP-medially for /s/, and 2.83 kHz IP-initially and 2.7 kHz IP-medially for /x/.[49] Repeated Measures ANOVAs indicate that there was a significant interaction between Accent and C_type ($F_{(1,9)}$=20.07, p=0.002), so the results will be discussed separately by the onset. The results of the four spectral moments of /s/ are shown in Figure 4.11.

Figure 4.11 Accent effects and spectral moments: fricative /s/

---
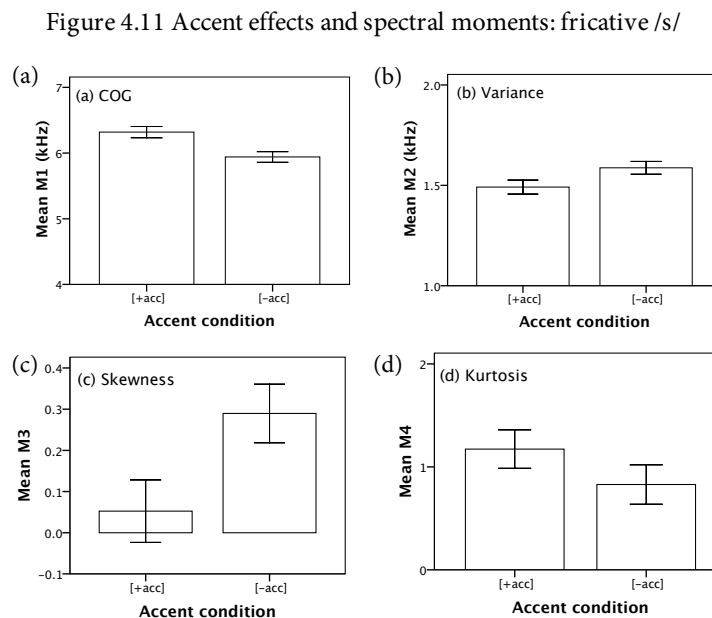
[49] Tests of Between-Subject effects show that Gender has a main effect on M1[$F_{(1,8)}$=12.86, p=0.007], but not on the others: M2 [$F_{(1,8)}$=3.6, p=0.094]; M3 [$F_{(1,8)}$=1.09, p=0.33]; M4 [$F_{(1,8)}$=0.01, p=0.99]. Since there is no interaction with Gender and the other factors, the results are reported pooled across gender.

The four figures show that accent effects are active in all four spectral descriptions. Repeated Measures ANOVAs show that there is a main effect of Accent on the first three moments (M1, M2, and M3), but not on M4: M1[$F_{(1,9)}$=32.86, p=0.000], M2 [$F_{(1,9)}$=8.88, p=0.015], M3[$F_{(1,9)}$=10.68, p=0.01], and M4[$F_{(1,9)}$=3.12, p=0.11]. This means that COG increases when /s/ is accented, and also the spectral energy distribution has less variance (M2) and is less skewed (M3) than the energy distribution of unaccented /s/. The difference in the fourth moment was not significant, but the trend shows that spectral energy is more diffused in the distribution.

The overall accent effect in /x/, on the other hand, is mostly absent except for M2 (variance). The /x/ results are shown in Figure 4.12.

Figure 4.12 Accent effects and spectral moments: fricative /x/
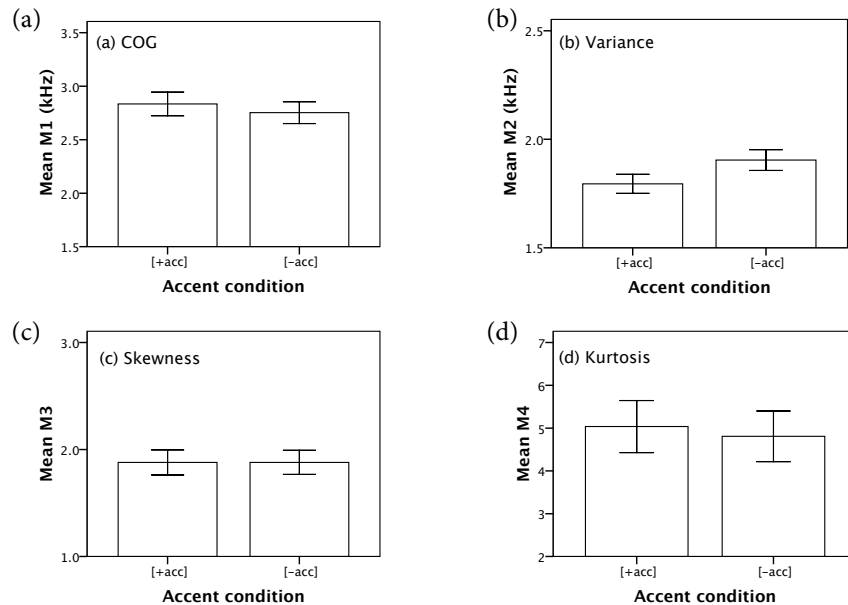
A higher COG is still associated with the accented /x/ as shown in (a), though statistically insignificant: M1 (p=0.4). Less variance, on the other hand, is found to be characteristic of the accented fricative /x/, suggesting that the second moment is the most consistent characteristic of the accent effect for the fricatives: M2 [$F_{(1,9)}$=15.64, p=0.003]. As we see in (c), skewness does not reflect any change in the accent conditions (p=0.67). The fourth moment shows the same trend

with the /s/ pattern such that kurtosis decreases in the unaccented condition. Although this decreasing kurtosis remains as a trend in both cases, the higher kurtosis in the accented fricatives could suggest that the high amplitude is more concentrated close to the mean frequency. In other words, the shape of spectra in the unaccented condition is relatively diffused relative to the accented one.

The spectral comparison made between /s/ and /x/ tells us that the accent effect is more visible in /s/ as measured by the four moments. One possible interpretation regarding the fricative identity and the spectral characteristics is that the prosodic manifestation in /s/ is more stable than /x/.

4.2.2 Interactions between accent and PoW effects

PoW effects have no main effect in all four spectral moments. However, the word level positional effect turns out to interact with the accent effect. The interaction between PoW and Accent is found in COG(M1), variance (M2), and skewness (M3), all of which are statistically significant: M1[$F_{(1,9)}$=8.55, p=0.019] and M2[$F_{(1,9)}$=9.36, p=0.016]. Since there is no interaction between PoW and consonant type, we focus on the /s/ results in most of the discussions.

In Figure 4.13, accent effects on COG are compared between the two positions in word.

Figure 4.13 PoW effect on COG and its interaction with accent: fricative /s/

104
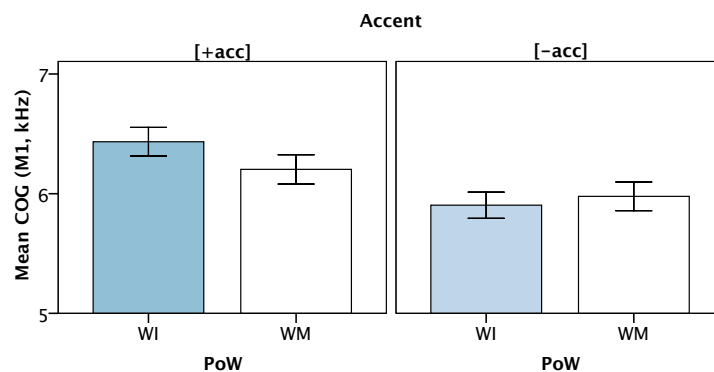
This figure shows how PoW effects interact with the accent effect. In the accented condition, the mean COG is higher word-initially than –medially, while the effect is absent in the unaccented condition. One-way ANOVA confirms that the COG mean is significantly different by PoW in the accented condition [$F_{(1,599)}$ =7.31, p=0.007], but not in the unaccented condition (p=0.36). The main effect of PoW on COG is marginally significant [$F_{(1,9)}$=5.18, p=0.52]. Recall that the accent effect was also to increase COG; hence, in the above figure, we find that the mean COG values of /s/ are higher in the accented condition regardless of PoW. The fact that the PoW effect is conditioned by the presence of accent suggests that the PoW effect may not be an independent prosodic effect; instead, it could be an enhanced accent effect in the word-initial position (at least in the spectral dimension).

The Variance (M2) result also shows a similar pattern regarding the PoW effect, as shown in Figure 4.14. That is, Variance in the spectral energy distribution is smaller word-initially, which was exactly the effect of accent as we saw in the previous section. The position-associated difference in Variance again disappears when unaccented.

Figure 4.14 PoW effects and the second moment (variance): result from fricative /s/
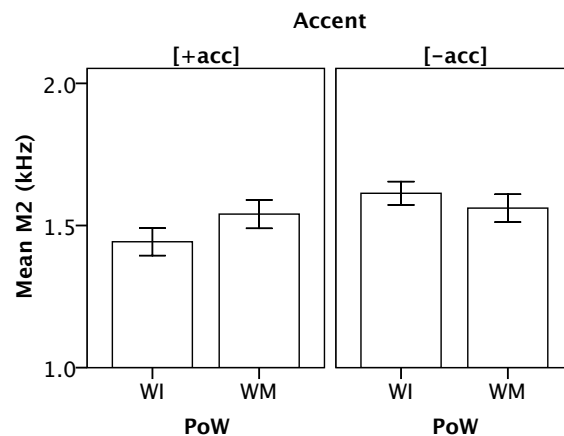
When /s/ is from accented syllables (on the left in Figure 4.14), variance is smaller word-initially. The difference in Variance (M2) by PoW is statistically different in the accented condition ($F_{(1,599)}$=7.84, p=0.005), but not in the unaccented ($F_{(1,599)}$=2.71, p=0.1) according to One-way ANOVAs. In addition, there is an interaction between Accent and PoW on the second moment

$[F_{(1,9)}=9.36, p=0.016]$. The result of the variance analysis by PoW confirms what was suggested in the first moment analysis: the spectral changes by PoW are attributable to the accent effect that is strengthened word-initially.

The results of the interaction between PoW and accent effects can be summarized as follows. Word-initial fricatives show higher COG (M1) and are less variable in their spectral energy distribution, but these characteristics are found only in the accented syllables. This suggests that the PoW effect in the spectral dimension can be viewed, for example, as enhanced accent effects that are sensitive to word position. As we will see in the following section, no interaction is found between PoW and stress. This means that the position sensitivity is specific to accent, at least in Spanish, since prosodic manifestations are assumed to be language-specific. One speculation is that word-initial position becomes prominent when the word is highlighted at the phrasal level (e.g., via accentuation), and thus, this positional sensitivity is irrelevant of the lexical level prominence.

4.2.3 Stress affects mainly the first spectral moment (COG)

Now we turn to the stress effect on spectral moments. There was a main effect of stress on COG (M1). A higher COG is found in the stressed fricatives when compared to the unstressed ones. The following figures show the overall stress effect on COG across the accent condition. As mentioned earlier, there was no interaction between Accent and Stress in the COG result.

Figure 4.15 Stress effects on COG (the first moment): fricatives /s/ and /x/
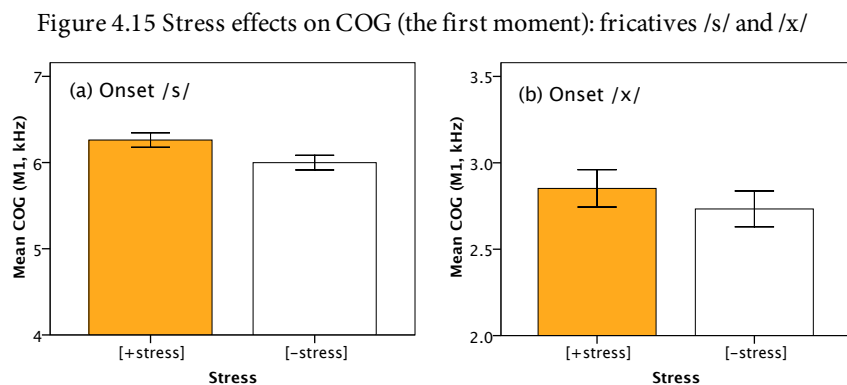
Stress effects on COG are shown in these figures for both /s/ and /x/. Overall, the main effect of stress is significant according to the Within-subjects analysis [$F_{(1,9)}$=11.87, p=0.009]. The effect seems to be weaker in the /x/ result, but there was no significant interaction between Stress and C_type. As mentioned earlier, the stress effect does not interact with the accent effect. Since both effects of stress and accent are independently present, the unstressed fricatives in the unaccented condition are expected to be the lowest in COG and the stressed ones in the accented condition are the highest. This prediction is confirmed in the data shown in Figure 4.16 below.

Figure 4.16 Stress and accent effects on COG: fricative /s/
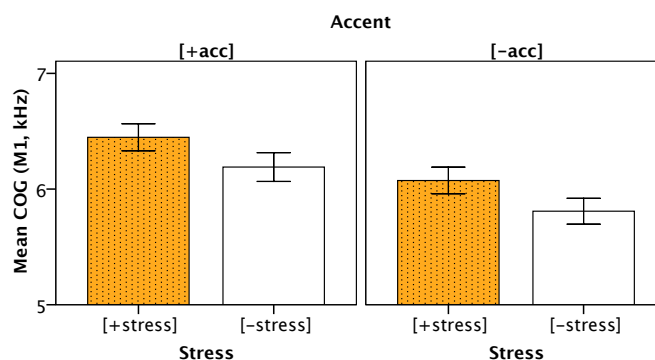
The figure shows that the common effect of stress and accent is to lower the mean frequencies of the fricatives. The mean difference between the stressed and the unstressed /s/ was even significantly different in the accented condition ($F_{(1,599)}$=9.07, p=0.003) according to One-way ANOVA. Thus, a higher COG (M1) characterizes stressed fricatives, and this finding is in line with the result for stops in the literature (Cho & Keating 2009). Other than M1, there was a weak effect of stress on M4 such that kurtosis increases in the unstressed fricatives: /s/ [$F_{(1,9)}$ =3.98, p=0.077] and /x/ [$F_{(1,9)}$=4.03, p=0.78]. This could be interpreted to mean that unstressed fricatives have a flatter spectral distribution indicated by heavier tails (higher kurtosis) in the spectral energy distribution.

4.3 Spectral balance as a reliable cue of stress in the vowel

Studies have suggested that changes in amplitude of particular frequency regions can influence the perception of fricatives, and these cues are stronger than the duration of fricative noise (Hedrick & Ohde 1993). This section focuses on such energy distribution over frequencies as a measure of spectral balance of the vowels. This analysis allows us to examine intensity changes over contiguous frequencies. Four filtered frequency bands are measured: B1 (0~0.5 kHz), B2 (0.5 ~ 1 kHz), B3 (1 ~ 2kHz), and B4 (2~4kHz) are measured, following Sluijter & van Heuven (1996b). These bands generally include the first four vowel formants. Under the assumption that stressed syllables would be produced with greater vowel effort, intensity at higher frequency bands is predicted to increase (Glave & Rietveld 1975) in the stressed vowels relative to the unstressed ones.

We first examine the spectral balance of vowel /a/, where the overall intensity of the vowel is also compared. As pointed out in Sluijter & van Heuven (1996b), vowel /a/ is the best place to find spectral balance differences as a function of stress, because F1 of /a/ is most fully separated from F0. Figure 4.17 shows the result when pooled across onset contexts since the two onsets do not significantly change the result.

Figure 4.17 Spectral balance of /a/ as a function of stress: IP-initial vs. IP-medial
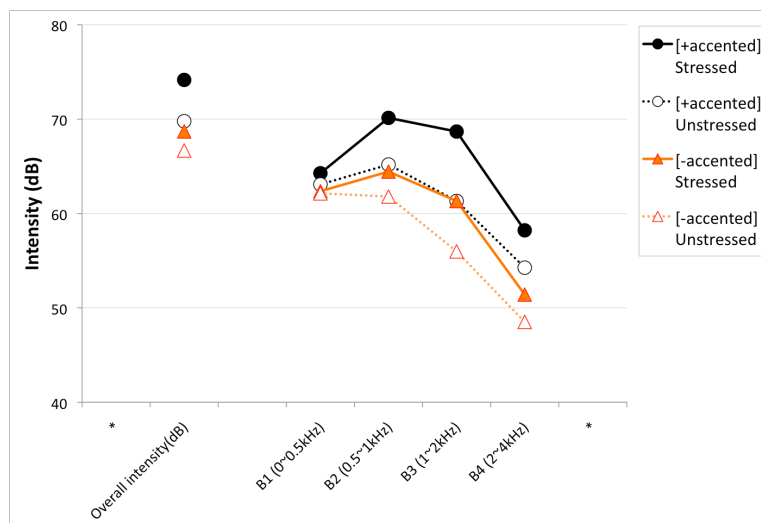
The overall intensity levels of vowel /a/ are different between the stress conditions within each accent condition. The difference was significant even in the unaccented condition: [$F_{(1,9)}$=9.53, p=0.013]. This indicates that stress effects are well characterized in our data by the differences in the overall intensity, even though the intensity differences were very small: 1.5 dB for [-accented] and 2.5 dB for [+accented] conditions. For the filtered bands (B1~B4), all bands except B1 (0~0.5kHz) differ in intensity as a function of stress, regardless of the accent condition [B1: $F_{(1,9)}$=0.17, p=0.694; B2: $F_{(1,9)}$=8.64, p=0.016; B3: $F_{(1,9)}$=48.95, p<0.001; B4: $F_{(1,9)}$=8.30, p=0.018]. The energy concentration in the figure appears at higher frequencies (B2~B3), and the spectral balance distinguishes the stressed /a/ from the unstressed /a/. The stress effect in the spectral balance is greater when the stressed vowels are also accented. These results are comparable to what has been reported in the literature (Sluijter & van Heuven 1996c; Sluijter et al. 1997).

The spectral balance in /a/ was examined first to be comparable to previous studies: for example, Sluijter & van Heuven (1996b), where /a/ was the most likely vowel if stress effects remain in the spectral balance in non-focal condition. They reported that the difference in the spectral balance as a function of stress disappears in vowel /ɔ/, if the vowel is located in a non-focal context. Our data also show variable stress effects on the spectral balance depending on the quality of vowel. The results are shown in Figure 4.18. The vowels are all taken from the IP medial position ([-accented]).[50]

---

[50] Repeated Measures ANOVA results are summarized in the Appendix (A- 6)

Figure 4.18 Spectral balance of the vowels as a function of stress: [-accented] position
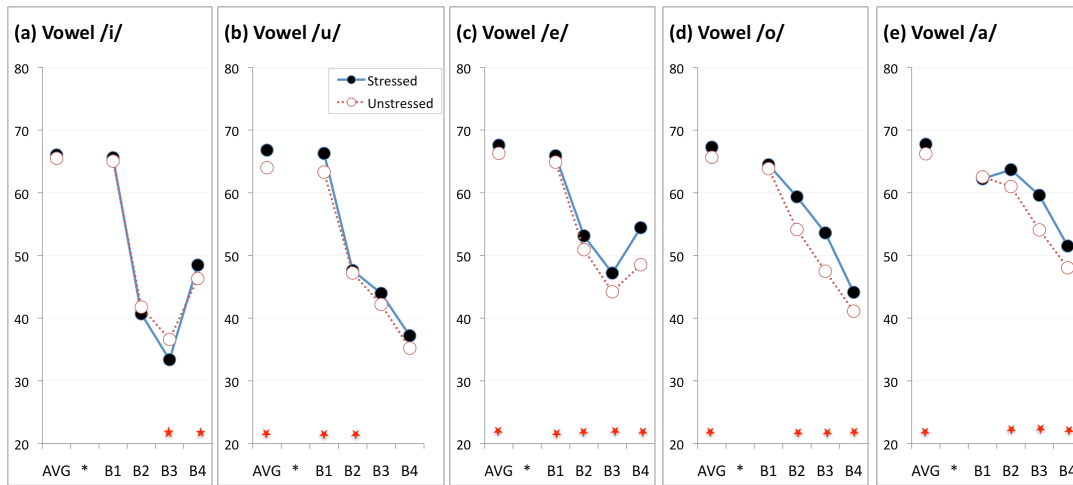
The stress effect in the spectral balance varies depending on the vowel quality. The high vowels have the smallest stress effect at the high frequency bands, and the non-high vowels /o/ and /a/ have the greatest. The Repeated Measures ANOVA results indicate that at least one of the high frequency bands significantly differs between the stressed and the unstressed vowels, as marked with an asterisk at the bottom of the figure. As for the overall intensity, all the vowels except /i/ can be distinguished by the stress conditions.

As mentioned earlier, the accent effect appears in the overall intensity, maintaining the stress effect in the spectral balance. The non-high vowels are compared between the accent conditions in Figure 4.19, illustrating the preserved stress effect and additional accent effects in the overall intensity.

110

Figure 4.19 Accent effects on the overall intensity: vowel /o/ and /a/

(a) Vowel /o/: [+acc] vs. [-acc]

(b) Vowel /a/: [+acc] vs. [-acc]

Legend:
- [+acc] Str
- [+acc] Unstr
- [-acc] Str
- [-acc] Unstr
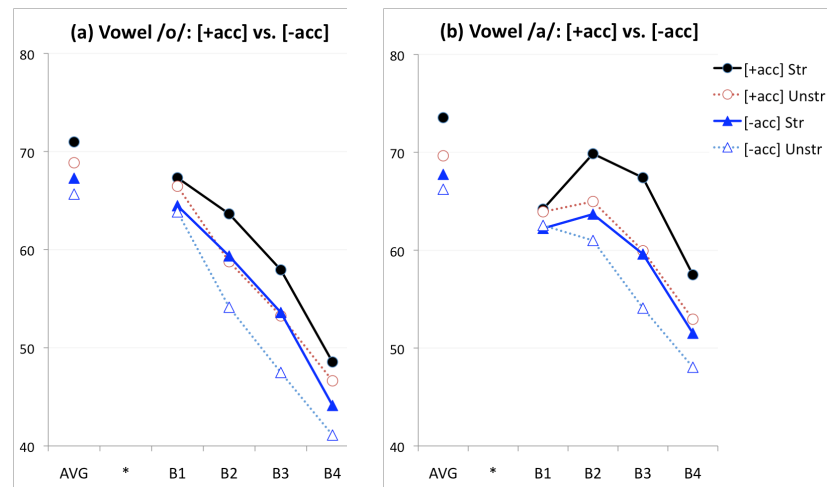
We see the accent effect in the figure with the increased overall intensity overall frequency bands. The stress effect within each accent condition is still well represented. This pattern is not limited to /o/ and /a/, and consistently found in the other vowels although the differences are smaller. This result has an important implication. In the figures, the spectral tilt over the bands overlaps when the vowel is either stressed or accented: compare the dotted line with empty circle to the solid line with triangle in Figure 4.19 (a) and (b). This means that stress and accent may influence spectral energy in the same manner, which implies that the stress and accent effects can be cumulative in influencing the spectral properties of vowel. We will develop this discussion in the summary section of this chapter.

The increasing intensity at high frequency bands was assumed to have a motivation of the potential increase in vocal effort in making stressed syllables more prominent. An interesting question arises regarding the PoW effect and whether the effect is to raise the overall intensity only, or it is similar to the stress effect that is visible in the spectral balance. If the PoW effect is similar to the stress effect, we expect the PoW effect present in the spectral balance: that is, the intensity increase will be greater word-initially and this would be independent of the stress conditions. A mixed result is found when we compare spectral balance results between /o/ and /a/ in Figure 4.20.

Figure 4.20 PoW effects in spectral balance: vowels /o/ and /a/
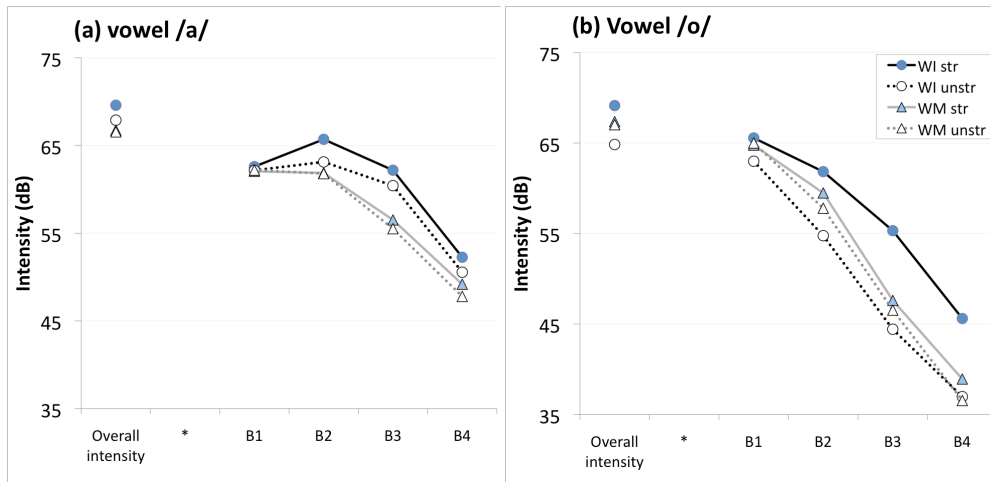
The intensity distribution for the vowel /a/ in (a) shows PoW effects in a similar direction to the stress effect. That is, we see that the word-initial /a/'s, regardless of the stress conditions, have greater intensity levels at the frequency bands B2-B4. In addition, the stress effect decreases word-medially, suggesting that the PoW effect may interact with the stress effect. The interaction seems to be stronger in the vowel /o/ as in (b), and the stressed /o/'s show greater intensity concentration over the high frequency bands regardless of PoW. What we can safely conclude from the two patterns is that the two types of prosodic effects can interact and influence the energy distribution at higher frequencies above 0.5 kHz. The energy at lower frequencies (0~5kHz) is hardly influenced by prosodic effects.

Now we consider how dialectal difference could be reflected in the intensity distribution. This examination can help us better understand the temporal result in Ch.3 by integrating the energy distribution result. Recall that the stressed vowels were longer in the steady-state (V-ss) for Argentinian compared to the Peruvian vowels. Our question is whether stress effects in the energy distribution also differentiate the two dialects as duration does. One-way ANOVA analysis shows the overall intensity (and B1) is significantly different between the two dialects. That is, the overall intensity (and B1) was higher in the Peruvian vowels compared to the Argentinian. This suggests that longer vowel duration (V-ss) does not necessarily accompany greater overall

intensity, nor vice versa. Figure 4.21 shows an example of the dialectal difference in the energy distribution. The data were taken from the unaccented /xa/ syllables.

Figure 4.21 Dialectal difference in the energy distribution: vowel /a/ in the /x/V syllables
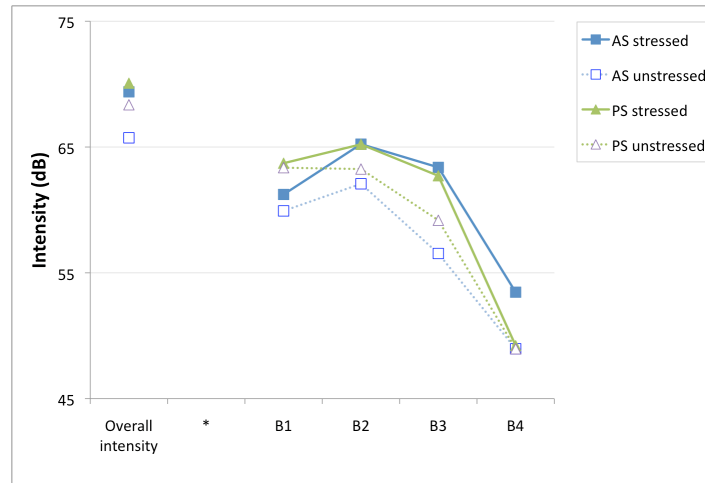
A noticeable difference is found in the overall intensity of the unstressed versus stressed vowels in Argentinian Spanish. This means that the intensity change is greater as a function of stress in this dialect relative to the Peruvian pattern. In addition, the stress effect in B3 is greater for the Argentinian vowel /a/ compared to the Peruvian, indicating that the intensity contrast by stress is greater for Argentinian at higher bands. Within-Subjects analysis shows that this dialectal difference is marginally significant: [$F_{(1,9)}$=5.185, p=0.052]. Thus, the unstressed vowels are weaker at higher frequency energy and shorter in duration for Argentinian when compared to Peruvian.

4.4 More peripheral vowels when stressed and/or accented

This section discusses the results of vowel formant measures to determine the effect of stress on vowel quality. Each vowel was preceded by either /s/ or /x/, and followed by a variety of consonant contexts. All target syllables were CV syllables placed word-initially and word-medially. For example, stressed /sa/ syllables appear in *sádico* word-initially and in *biságra* word-

medially. A total of 2400 tokens were included in the formant analysis for each vowel: 2 Dialects × 2 Accent conditions × 2 Stress conditions × 2 Fricatives in onset × 5 vowels × 2 PoW × 5 Subjects × 3 Reps = 2400 tokens.

Vowel formants were measured at the mid point of the vowel steady-state (labeled V-ss), in order to minimize contextual effects from neighboring consonants and vowels. The speech data were sampled at 44kHz and analyzed in Praat (ver. 5.1.31) to obtain the first three formants. Formant tracking was implemented in a script with parameters set to 10 ms fixed time step with a 25 ms window for spectral analysis. Additional adjustment and corrections were performed manually for outliers in the formant results.[51]

The influence of stress on vowel quality is abundantly evident in the grammar of spoken languages. For instance, many stress system languages show some degree of restriction on phonemic vowel contrasts depending on the presence or absence of stress. In general, reduced vowels can be found in unstressed syllables, while full vowels appear in stressed syllables and/or are only selectively allowed in unstressed syllables (Partridge 1950). In addition, a number of studies have explored the acoustic correlates of stress and concluded that vowel quality is affected by stress as well as other phonetic factors (Bolinger 1958; Lieberman 1960; Beckman 1986; Sluijter & van Heuven 1996c). For example, stressed syllables are associated with more peripheral vowels than unstressed syllables, which in turn tend to be associated with less peripheral and centralized/neutralized vowels (Jones 1956; Tiffany 1959; Lindblom 1963; Delattre 1969; Gay 1978; Fourakis 1991; Cho 2005).

As for stress effects on the vowel quality in Spanish, it has been described that vowel reduction is rare (Hualde 2005). This is evident from the fact that studies examining stress effects in Spanish do not consider vowel quality as a parameter. Although stress effects on vowel quality in Spanish may not be as evident as in other languages such as English, Dutch, and Swedish, we find descriptions suggesting that vowel quality in Spanish is also influenced by stress. For example, it has been reported that speakers tend to reduce the F1-F2 distance in unstressed [e]

---

[51] Outliers were statistically defined: the tokens outside ± 2 SD were manually inspected, and corrected if identified as measurement errors.

(Ortega-Llebaria & Prieto 2005; Ortega-Llebaria 2008). The effect, however, decreases in post-focal sentences and disappears in focal contexts. This means that vowel quality in Spanish is sensitive to prosodic factors such as accent and stress. In addition, we have seen in Chapter 3 that the vowel durations were significantly different between stressed and unstressed conditions. If vowel reduction depends entirely or partly on the duration of vowel (e.g., Lindblom, 1963; Sluijter & van Heuven 1996), we would expect stress and/or accent effects in the vowel formant to some extent. This chapter examines prosodic effects on the vowel quality as a measure of the first two formants.

Formant measures of the five phonemic vowels are summarized in Table 4.1 for female (N=4) and male (N=6) speakers.

Table 4.1 Vowel formant mean summary

	Male		Female	
V	F1 (SD)	F2 (SD)	F1 (SD)	F2 (SD)
/i/	319 (39)	2193 (225)	374 (45)	2643 (259)
/u/	365 (43)	1011 (231)	427 (39)	1128 (247)
/e/	420 (58)	1850 (170)	529 (94)	2277 (199)
/o/	458 (69)	1049 (181)	573 (86)	1217 (222)
/a/	646 (94)	1447 (118)	857 (109)	1782 (133)

(Unit in Hz)

Gender effects in the two vowel formants were significant: [$F_{(1,9)}$=17.76, p=0.002]. One-way ANOVAs show that all the five vowels are distinguished by gender.[52]

As is well-known, gender effects have a physiological source (e.g., the size and length of the vocal tract/pharyngeal cavity among others), which involves formant frequencies in such a way that lower resonances are produced through a larger/longer vocal and/or pharyngeal cavity,

---

[52] A complete table for the statistical result appears in the Appendix.

all else being equal (Fant 1966; Johnson 2005).[53] Thus, when we need to confirm that vowel results are consistent between genders, we will make use of vowel normalization process, through which results from different gender could be comparable enough to interpret common patterns out of the results.[54] The following figure (a) is the vowel formant results from male and female, and (b) is the normalized vowel space between genders. The summary of the results was shown earlier in Table 4.1.

Figure 4.22 Vowel formants pooled across conditions: raw data and normalized vowel space

(a) Vowel space by gender in Hz   (b) Normalized vowel space: Labov method

---

[53] The size of vocal tract/pharyngeal tract differs approximately by 20% between genders (Simpson, 2001), but, for example, F0 differences between genders are not proportional to this number. That is, F0 of male speakers (i.e., 120 Hz) could be a half of the females' (i.e., 210 Hz) on the average.

[54] Labov's method is the normalization approach used by the phonological 'Atlas of North American English' (see the details of the method at http://ncslaap.lib.ncsu.edu/tools/norm/). This method is a modification of Nearey (1977)'s normalization using a log-mean for formant normalization. It is a speaker-extrinsic normalization method and is adopted here to normalize gender difference. Speaker-extrinsic methods are designed to normalize varying vowel formant values that supposedly originate from different sizes of vocal tract shapes. Therefore the vowels produced by speakers within a dialect are normalized using this method in this study. In addition, there are other aspects of gender effect on speech production that cannot be accounted for solely by anatomical differences. For example, acoustic vowel space cannot be proportionally scaled based on any physiologically determined factors (Fant 1975; Simpson 2001; Peterson & Barney 2002; Johnson 2005). The current study cannot discuss this socio-cultural factor and leaves it for future research.

The vowel space on the left is from the raw data shown in Table 4.1. The gender difference in the vowel space is distinct, although the overall distributions are quite similar between the two. The results were put to NORM normalization process, and shown in (b). The normalized vowel spaces become similar in scale after the normalization, except the vowels /u/ and /a/. We conclude that the normalization process can help us understand consistent patterns across gender and will be used in the following discussions when we show across-gender results.

We turn to accent and stress effects on the vowel distribution. Repeated Measures ANOVAs were conducted, and show that there were significant main effects of Stress and Accent, and also an interaction between Accent and Stress. As a main effect, accent increases F1 value of vowels. This effect was statistically significant [$F_{(1, 9)}$=19.70, p=0.002]. The increasing F1 can be interpreted to mean that vowels are articulated more open in the accented condition than in the unaccented condition. F2 also tends to increase for the front vowels when accented, but this pattern was not statistically significant [$F_{(1, 9)}$=2.28, p=0.17]. Figure 4.23 exemplifies the accent effect: the vowel [a] is considered to show the accent effect on F1 and the vowel [e] for F2. These are shown in (a) and (b), respectively.

Figure 4.23 Accent and stress effects on the F1 and F2: vowel /a/ for F1 and /e/ for F2

(a) F1: Vowel [a]   (b) F2: Vowel [e]
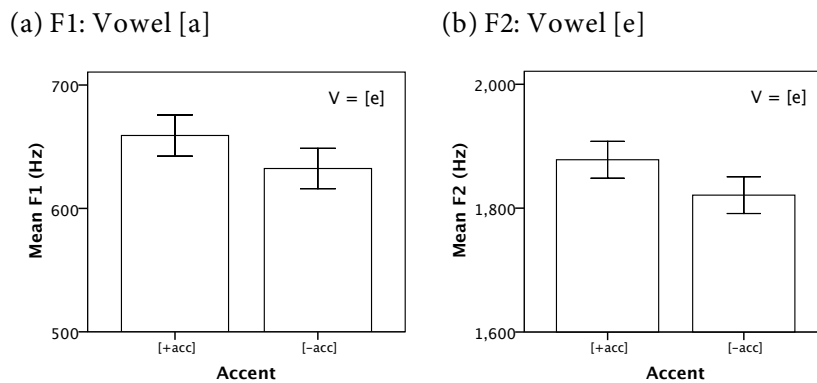
We see that accent influences F1 and F2 such that the vowel formants tend to be higher in the accented condition. However, as mentioned earlier, the accent effect on F1 is significant, but not on F2.[55]

The interaction between Accent and Stress is such that the stress effect is greater when the stressed vowels are also accented. This interaction predicts that the stressed vowel formants are the most peripheral in the vowel space, while the unstressed ones are the most centralized in the vowel distribution. Stress as a main effect was also significant. Stressed and unstressed vowels are grouped separately and the vowel space of each group is shown in Figure 4.24.

Figure 4.24 Accent effects on the vowel formants: results from the male vowels

(a) Stress effects in the accented condition   (b) Stress effects in the unaccented condition
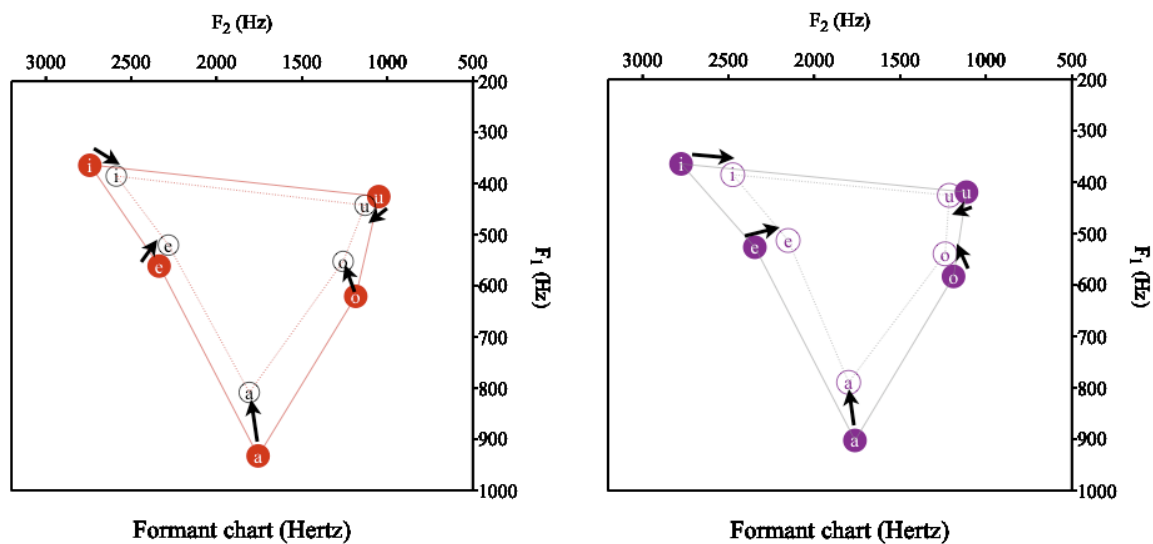
Similar to the accent effect, F1 is significantly influenced by stress even in the unaccented condition (b). Stress also affects F2 values, though this is statistically not significant. This stress effect is consistent with the accent effect in the F2 dimension.

---

[55] There was a main effect of C on both F1 and F2 [F1: $F_{(1,9)}=40.35$, $p<0.001$; F2: $F_{(1,9)}=93.96$, $p<0.001$]. When the vowel follows onset /x/, F1 is higher compared to the vowels following /s/. This pattern is significant only for the non-high vowels. C effect on F2 is to increase the values in the accented condition, and this effect is greater in the front vowels.

We have seen so far how the vowel formants are influenced by stress and accent. The results indicate that the two prosodic effects are similar in that the vowel F1 increases when stressed and/or accented. In other words, vowels are articulated as more open when stressed and/or accented.

The vowel formant results are normalized to determine whether the stress and accent effects are consistent between genders. The normalized vowel formants between genders are plotted in F1 and F2 by the accent conditions: (a) accented and (b) unaccented.

Figure 4.25 Stress and accent effects in the normalized vowel space

(a) Accented condition

(b) Unaccented condition
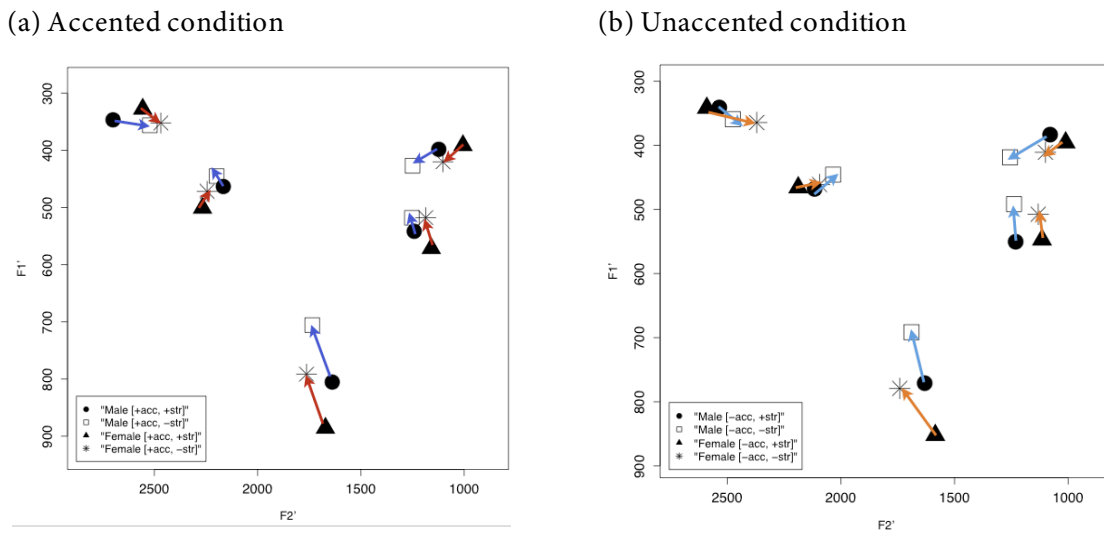
Let us examine first the accented condition in (a). In the normalized vowel space, the stress effect is represented in a consistent way. For example, the vowel [a] between genders show almost identical changes by stress. The stress effect within each vowel category shows a similar pattern between genders in the direction and degree of changes in F1 and F2. The unaccented condition in (b) also confirms that the stress and accent effects are consistently found in the vowels of different genders, and the direction of the effect in the F1 and F2 is identical. Thus, we conclude that the stress effect and accent effect are consistently found regardless of gender and the direction of the effects are similar such that unstressed and/or unaccented vowels are more centralized in the vowel space, particularly in F1 dimension (openness).

119

Our final examination is to compare the stress effect on vowel quality between the two dialects. This is to investigate whether the temporal variation between the two dialects is linked to spectral changes in the vowel. It will also enable us to interpret the overall relation between temporal and spectral modifications motivated by prosodic factors. The mid vowel is chosen based on the assumption that the mid vowel may have a relatively higher degree of freedom than the other corner vowels (e.g., articulatory ceiling/floor effects for the other vowels). There appeared one characteristic behavior of the non-low front vowel in the vowel space of the female Argentinian Spanish. That is, the mid vowels /e, o/ in AS were raised in F1 dimension when they are stressed, leaving more space between /e, o/ and /a/. The vowel formant results from females are shown in Figure 4.26, for the two dialects to be compared.

Figure 4.26 Dialectal comparison of vowel space by stress: female vowels

(a) Argentinian Sp. (female vowel space)   (b) Peruvian Sp. (female vowel space)
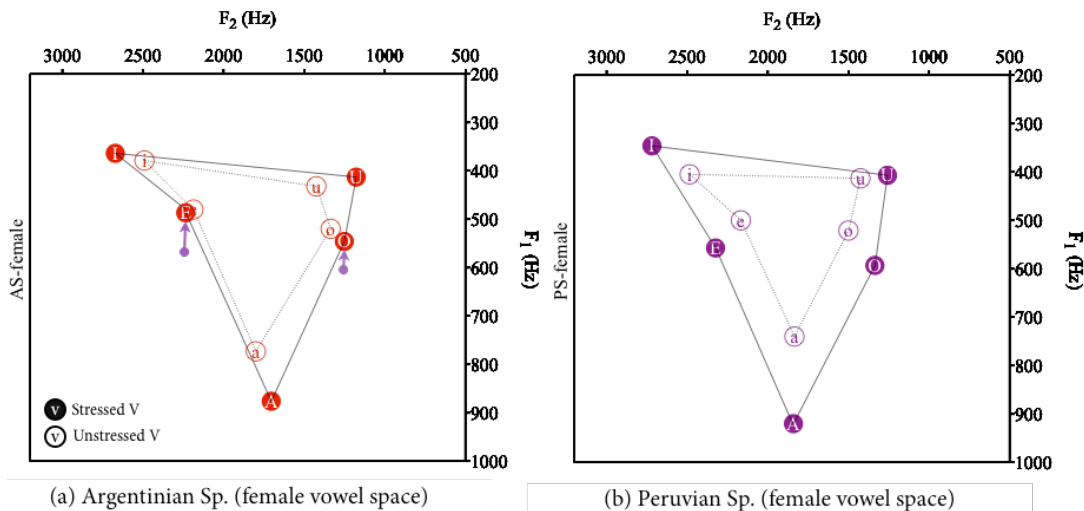
The vowel space of the two dialects are compared in (a) and (b). In Figure (a), the two mid vowels /e, o/ are closer to their high vowel counterparts in comparison to those in PS (b). The arrows in (a) represent the mid vowel difference between the two dialects. The difference in the mid vowels is clearly for the F1 of the stressed vowels in AS. Thus, the raised F1 is attributable to the stress effect on the vowel in AS. This formant change itself can have variable interpretations acoustically and articulatorily. To have relevant interpretations to our interest, we combine the results of vowel duration and spectral analysis. Recall that the stressed vowel steady-state (V-ss)

was longer and the spectral energy at higher frequencies was greater in AS than in PS. The results are shown in Figure 4.27.

Figure 4.27 Dialectal comparison of stress effects in duration and spectral energy distribution

(a) Stress effect on vowel Duration
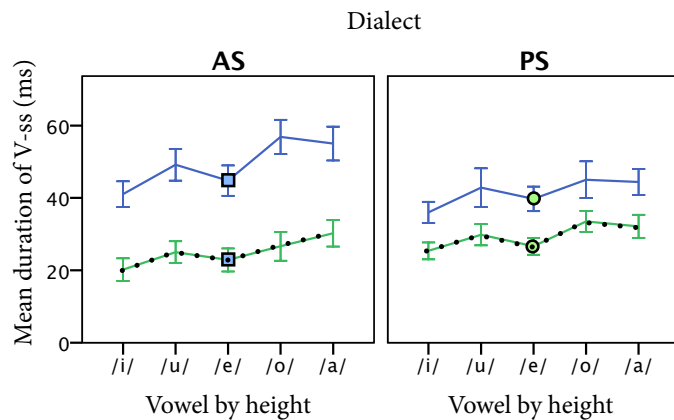
(b) Stress effect on spectral energy distribution
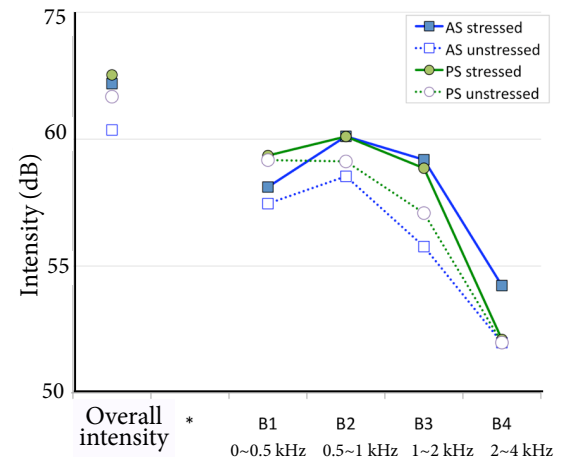
This means that the F1 raising of the mid vowel in AS is not due to the undershoot triggered by temporal reduction (Lindblom 1963; Fourakis 1991). In addition, the spectral energy distribution in (b) indicates that the stress effect was greater in AS than in PS, as the energy changes above 0.5kHz (B2~B4) are greater in AS (Sluijter & van Heuven 1996a). We have two possible interpretations. First possible interpretation is that the decrease in F1 for the stressed mid vowel represents a greater CV coarticulation in AS. Stress as well as accent is known to influence CV coarticulation such that stressed (also accented) segments often resist coarticulation (De Jong et al. 1993; Magen 1997; Lindblom et al. 2007). Given the fact that CV coarticulation is weaker in the stressed syllables (e.g., no pressure from temporal reduction), the dialectal difference in the mid vowel F1 can be interpreted to mean that the coarticulation of CV (/s/V or /x/V) was less resistant in AS than in PS. It is known that vowel raising of mid vowels can occur more frequently after a preceding anterior (or palatal) consonant: e.g., mid-vowel raising in Puerto

Rican Spanish (Navarro 1948; Holmquist 1998).[56] Alternative interpretation is that the mid vowel in AS reflects hyper-articulation under stress (Lindblom 1990; De Jong 1995). It is known that hyperarticulation is speaker-specific and may have different strategies for different segments. That is, for low vowel, speakers often show more jaw opening, but they show tongue dorsum raisin for high vowel (De Jong 1995). The variable strategies are often interpreted to mean that paradigmatic contrasts (or contrastive features) are enhanced under stress effects. If this is the case for the mid vowel behavior in AS, the hyperarticulation hypothesis suggests that the mid vowel enhances a non-low feature (rather than non-high feature) and thus raises F1 under stress. Whether the dialectal difference is accounted by the difference in CV-coarticulation effects or the hyperarticulation strategy, we can see that the stress sensitivity of vowel is different between the two dialects.

In addition, the back vowels /o, u/ include effects of rounding, which is makes our comparison more complicated. We will examine stress and accent effects on F1 values of [e] between the two dialects. The main effect of accent is present in the two dialects as well as stress effects. There is a three-way interaction among Accent, Stress, and Dialect on F1: $[F_{(1, 9)}=5.57, p=0.046]$. To interpret this interaction, we compare the interaction between Stress and Accent in each dialect. The means of F1 values of [e] will illustrate the dialectal difference in Figure 4.28.

Figure 4.28 Stress×Accent interaction in the F1 of [e]: a dialectal comparison

---
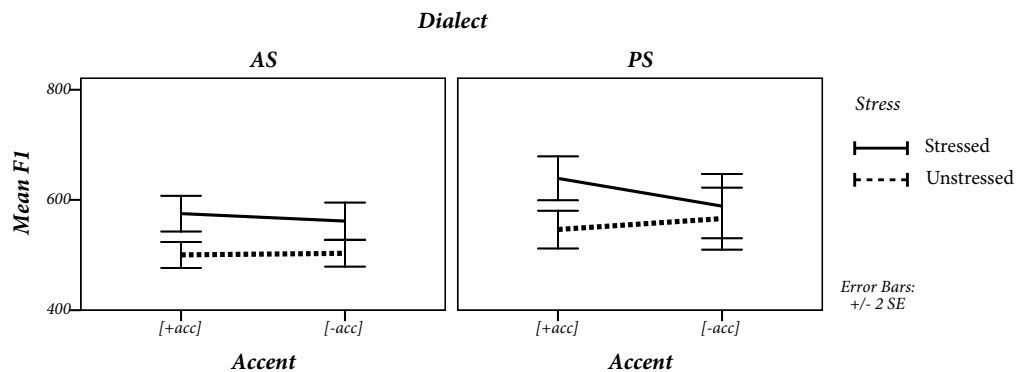
[56] Mid-vowel raising has been found in other peninsular Spanish dialects (McCarthy 1984; Holmquist 1985; Hualde 1989), and also in other Romance languages (e.g., Italian dialects: Maiden 1991), although the sources of the vowel raising are diverse.

In the figure, we see that stress effects in the unaccented condition disappear in the Peruvian [e], but are preserved in the Argentinian [e]. This can be interpreted to mean that the vowel [e] does not change in F1 in the unaccented condition. If we interpret the decrease of F1 as a sign of reduced vowel quality, the insensitivity of [e] to the stress effect can be interpreted to mean that the Peruvian vowels are not reduced in the unaccented condition. That is, the vowel [e] in Peruvian Spanish preserves the vowel quality regardless of stress if the vowel is unaccented. Recall that the duration of the vowel steady-state was shorter in general for Peruvian Spanish than Argentinian Spanish (Table 3.5). If vowel quality is conditioned by temporal variation, the result of this vowel formant pattern in [e] is not surprising. That is, the shorter vowel steady-state in the unstressed vowels in Argentinian Spanish might have induced an incomplete vowel target (e.g., undershoot).[57]

4.5 Summary of stress effects in the spectral and intensity dimensions

This chapter has focused on how spectral energy distribution in the onset fricatives and the vowels is influenced by stress. First, the energy level of each syllable constituent was measured in RMS intensity, and the overall intensity in dB when we discussed spectral balance in the vowel. Secondly, spectral energy distributions were investigated as measures of spectral moments for the fricatives and a measure of spectral balance and formant distributions for the vowels. Each measure of spectral energy and its distribution has revealed important effects of stress besides other prosodic effects including accent and PoW.

---

[57] The other vowels are also examined individually to see how consistent this pattern is across the vowel categories. The results show that the formant change in F1 is smaller in Peruvian only for the mid vowels [e] and [o]. The low vowel [a] shows a similar degree of change in F1 between the two dialects. The high vowels, on the other hand, show the opposite pattern such that the F1 vowel formants of the high vowels are more sensitive to stress in Peruvian Spanish.

First, the RMS result showed that the stress effect as well as the accent effect mainly influences the vowel energy. The stressed vowels have greater RMS than the unstressed vowels, and this effect is present regardless of the presence and absence of accent. In addition the stress effect is greater when the stressed vowels are accented. This means that stress and accent effects are cumulative. This cumulative characteristic can account for the RMS difference in the onset fricatives when they are accented. That is, the consonant RMS differed between stressed and unstressed fricatives when they are in the accented position (see Figure 4.1), although stress and accent did not have a main effect individually on the consonant RMS. We also investigated the stress effect on the energy level by the RMS C/V ratio based on the assumption that relative strength in the energy could be more important in determining stress/accent effects on the syllable as a domain. The result showed that stress effects could be well characterized by the increase in the RMS C/V ratio (see Figure 4.7) across the accent conditions and the onset types. One may argue that the changes of the vowel energy would be enough to characterize stress effects, since the increasing C/V ratios in the unstressed syllables were mainly due to the decrease of the energy in the vowels. One advantage of the C/V ratio analysis in the current study was that we could compare the stress effect on the syllable in general between the two dialects. That is, the C/V ratio change as a function of stress was greater in the Peruvian syllables compared to the Argentinian, even though the RMS energy for the consonant and the vowel was overall higher in Peruvian. One interpretation regarding this dialectal difference was that the unstressed onsets in the Peruvian syllables were not as weak as Argentinian ones. In other words, the Peruvian onset fricatives retained its energy level when they are unstressed and at the same time unaccented.[58]

Secondly, spectral energy distribution in the onset fricatives was examined as measured by the four spectral moments. The second moment was the most reliable parameter to describe the stress effect on the spectral distribution of the two fricatives. For the alveolar fricative, the first three moments could all characterize the stress effect, while only the second moment (Variance of the spectral energy distribution), was available for the velar fricative. The COG was higher in the stressed and/or accented /s/, and this remained as a trend for /x/. The Variance analysis could

___

[58] Alternatively, it could be that the Peruvian fricatives are not sensitive to stress and/or accent.

describe stress and accent effects such that the spectral energy distribution is less variable when the fricatives are either stressed or accented, as it means that the deviations from the mean spectral energy over the frequencies is smaller.

The third analysis focused on the spectral energy distribution in the vowel. The energy changes over frequencies were investigated by comparing the four filtered frequency bands. The prediction from the literature was that only the higher band energy would increase as a function of stress, and our results confirmed that the spectral balance could be a reliable cue to the stress effect. In addition, accent showed the same effect: higher energy increased more when the stressed vowels are also accented. Our results showed that the spectral balance difference was robust and not limited to the low vowel [a]. To a lesser degree, the other vowels were also characterized by this parameter (see Figure 4.18). The interaction between stress and accent effects showed that the two effects could be cumulative. That is, energy at higher frequencies increased when the vowels are either stressed or accented, and it was greater when the vowels are stressed and at the same time accented.

Stress effects in the vowel were also investigated by measuring the first two vowel formants. The result showed that the vowels were mainly affected by stress in the F1 space (openness), and weakly in the F2 (front-back tongue position). That is, the stressed vowels increase in F1 (and optionally F2). There was also a main effect of accent in the vowel formants in the same direction with the stress effect. This means that vowels are more peripheral when they are stressed, and even more when they are also accented. On the other hand, the unstressed vowels changed in the two-dimensional vowel space (more in F1 and optionally in F2), resulting in a relatively narrower vowel space. This result confirms the finding of Ortega-Llebaria & Prieto (2006, 2011). Thus, the vowel quality in Spanish is also affected by stress, and this stress effect may increase in the presence of accent.

The multiple analyses on the spectral characteristics of the onset fricatives and the vowels have confirmed the robust effects of stress and accent. In addition, we found PoW effects specific to the onset fricatives. Contrary to the PoW effect in the temporal dimension (Ch.3), which was to lengthen the duration of word-initial fricatives, the intensity of fricatives was greater word-medially than –initially. We interpreted this result such that the PoW effect is mainly active in

the temporal modulation rather than in the energy level. The spectral analysis of the fricatives as measured by the spectral moments showed that the COG of the word-initially located fricatives is higher than their medial counterparts. However, this pattern was present only when the fricatives were accented. This conditional effect further suggested that the PoW effect might not be an independent prosodic effect at least in the spectral dimensions; instead, it could be an enhancement of accent effects in the word-initial position. Based on the intensity and spectral results, we concluded that PoW is involved primarily in the temporal modulation of onsets.

To summarize, the investigation of the spectral characteristics has provided evidence that the spectral modification in the syllables occur as a function of stress even when the syllables are unaccented. As pointed out in Sluijter & van Heuven (1996), the insight of the spectral modifications for the stressed syllables, both smaller C/V ratio and the greater energy at higher frequencies, can be best understood to influence the perceptual loudness of stressed syllables, which rely more on the higher frequency energy.[59]

---

[59] Sluijter & van Heuven (1996a) refers to Zwicker & Feldtkeller (1967) for this interpretation: "…(the study) showed that the energies in the low-frequency bands add little to perceived loudness, while the contribution of the higher bands is much stronger."

Chapter 5 Discussion and conclusion

This dissertation investigated the phonetic realizations of syllables influenced by prosodic factors including stress, accent, and position in word. Our investigations of prosodic factor-specific manifestation have shown that stress effects are extensive in duration and spectral properties, independently of the other prosodic factors, and the effect can differentiate stressed onsets as well as vowels from their unstressed counterparts. Prosodic interactions were also investigated among stress, accent, and position in word effect. Variable interaction patterns were found in a number of phonetic properties which combine in complex ways that are likely language-specific. We attributed the apparent language-specific patterns of prosodic interaction to a higher-level rhythmic property of Spanish (e.g., isochrony), which defines an upper bound in constraining lower-level prosody manifestation. Our main results are reviewed in this chapter and are integrated to address the research questions set out in the Introduction (Chapter1).

The first research question was what overall acoustic modifications are produced by prosodic factors, and specifically, how phonetic realizations of onset and vowel are affected by prosodic effects. We compared prosodic effects in duration, intensity, and spectral characteristics of C and V (measured by COG and spectral balance). The results showed that duration and spectral properties could differentiate prosodically strong segments from their weak counterparts. Although there were more consistent prosodic effects on vowels, onset consonants were also affected by prosodic effects. These prosodic effects in relation to syllable constituents are discussed in Section 5.1.

Our second and the third research questions are discussed in Section 5.2, which addresses whether stress manifestation is independent of other prosodic effects, and how the three prosodic factors interact. Our results show that stress has its own phonetic manifestation, and it interacts with other prosodic factors in specific and independent ways. This section will summarize what we have learned about the phonetic nature of each prosodic effect. We also aimed to disentangle stress effects from other prosodic effects, focusing on phrasal level of prominence (accent) and

position in word effects. The results showed that stress and accent are similar in the ways they affect phonetic properties of segments, but different in the ways they interact on specific phonetic dimensions. The stress and accent effects could also be distinguished from the position in word effect, which was constrained by the accent effect. The similarities in the behavior of stress and accent effects imply that stress and accent together behave as one kind of prosodic effect (e.g., prominence effects), and position in word as a different kind (e.g., boundary effects at low-level). This prosodic factor comparison will answer the third question of how prosodic factors are interrelated in the phonetic manifestation of segments. In Section 5.2, we will discuss how cumulative strengthening effects of prosody can be predicted when multiple prosodic factors are concomitant in syllables.

In Section 5.3, we consider isochronous rhythm as an upper bounding prosodic constraint on the syllable internal variation found in our results. When we compared sub-segmental temporal modifications within syllable constituents, we found that the two dialects differed in the way stress affected the durations of the sub-phonemic components of the syllable onset and the vowel. The asymmetric pattern found in the two dialects calls for an account of within-syllable temporal modifications under stress. A gestural account is provided, as implemented with a computational simulation in TADA (TAsk Dynamics Applications). This gestural representation of prosody manifestation can accommodate syllable internal temporal variation.

5.1 Acoustic modifications of stressed syllables in relation to syllable constituency

We compared durational and intensity changes by stress with special attention to syllable constituency. The results can be summarized as follows.

The durational difference between stressed and unstressed syllables was found to be significant in sub-segmental units (transitions and steady-states). Even though Spanish has been described as preserving phonemic quality in unstressed syllables, stressed vowels were differentiated from unstressed vowels in quality measured by F1 and F2. Spanish indeed shows a relatively small durational difference by stress, but the stress effect on duration of syllables is

consistent.[60] Thus, temporal modulation by stress is clearly regulated by a language-specific mechanism, and it cannot simply be described as a mechanical articulatory effect which predicts no cross-linguistic differences. RMS intensity was also compared between stressed and unstressed syllables, and the results showed that stress effects in the intensity dimension are mainly on the vowel, and are not significant on the onset. Thus, the overall intensity difference between stressed and unstressed syllables can be attributed to prominence effects on the vowel.

In the following subsections, we discuss how stress effects on the rhyme are related to those on the onset by comparing stress effects on each constituent in two dialects of Spanish.

5.1.1 Stress effects on onset

The acoustic properties of onsets were assessed via duration, RMS intensity, and spectral measurements. We have seen that onsets are greatly affected by stress in a number of phonetic properties. Differences between the two dialects, which suggest that there could be an onset-rhyme interaction in the phonetic manifestation of stress, were found in the duration and spectral results of vowels (mid vowel raising).

The duration results were based on two measurement methods. The first followed a traditional way of marking the boundary (i.e., voicing initiation) between consonant and vowel within the syllable. The second measurement, the narrow method, divided each syllable constituent further into sub-segmental units such as transitions and steady-state parts. In the broad method, the two dialects were almost identical in the durations of C and V and the way that the syllable duration was affected by stress. Although the duration of onset fricatives was slightly longer for PS than for AS, and vice versa in the duration of the vowel, this pattern remained only a tendency. On the other hand, the narrow measurement results showed that stressed syllables modulate their subparts within the syllable constituent, and that dialects can differ in how this modulation occurs. We saw that the sub-segmental patterns are distinct

---

[60] Stressed/unstressed syllable duration ratio was 1.28 in Spanish, and this is smaller compared to English 1.8 or 2 (Parmenter & Treviño 1935).

between the two dialects: significantly longer onset steady-states (PS) or longer vowel steady-states (AS) in stressed syllable. Figure 5.1 (repeated from Figure 3.11) represents the durational composition of each subunit, and demonstrates the asymmetric stress effect between the two dialects.

Figure 5.1 Sub-phonemic durational changes by stress: a dialectal comparison
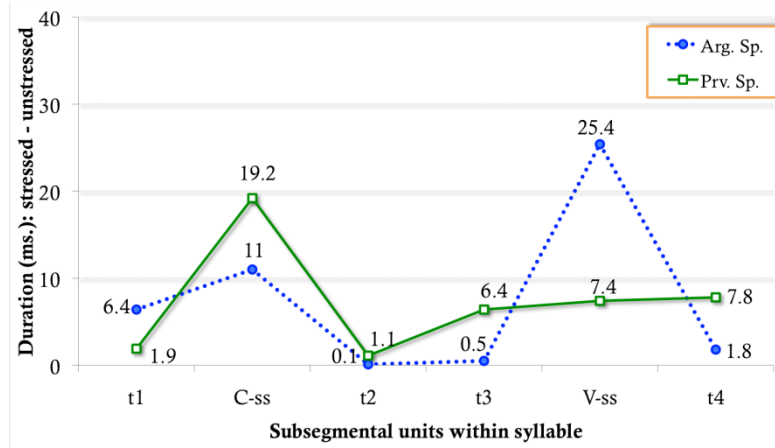
Recall that the durations of the unstressed syllables were almost identical in the two dialects. When syllables are stressed, a greater durational change is found within the onset (C-ss) for PS while it is in the vowel (V-ss) for AS. This pattern suggests that intra-syllabic effects of stress can be variable. This result is in line with the articulatory results of Cho & Keating (2007), which showed stress effects on the articulatory seal duration with no overall durational changes in the onset. This can be interpreted to mean that stress effects in the articulatory magnitude can vary cross-dialectally (and cross-linguistically) even without apparent durational differences when measured by abstract phonemic units (such as C and V). This result also highlights an advantage of the narrow measurement method used in this study, which can be further exploited in future studies.

The spectral investigation provided a clearer idea of how the two dialects differ in the magnitude of stress effects. The overall COG values of the onset fricatives were greatly influenced by stress, and the effect was greater in PS than in AS. This was interpreted to mean that the

spectral quality (when measured by COG) as well as the sub-segmental duration of the onset fricatives is more sensitive to stress in Peruvian Spanish than in Argentinian Spanish.

5.1.2 Stress effects on rhyme vowel

Spectral balance and vowel formants were used to characterize the quality changes in the vowel when stressed. It was found that stress increased spectral energy, mainly at frequencies above 500 Hz. Stress effects in the spectral energy distribution of vowels were significant, and were different between the two dialects, with energy change at higher frequencies being greater in AS (Figure 4.21). In the vowel formant analysis, stressed vowels were more peripheral compared to unstressed vowels, in both dialects. This vowel space expansion or reduction depending on stress conditions is well acknowledged in the literature (Tiffany 1959; Lindblom 1963; Fourakis 1991), and Spanish vowels also undergo quality changes affected by stress, although to a lesser degree.[61]

We have seen that the mid vowels /e, o/ in AS were raised in F1 when they were stressed, leaving more space between the mid vowels (/e, o/) and the low vowel (/a/) (see Figure 4.27). When we combined the results of vowel duration and spectral properties, the stressed vowel in Argentinian Spanish was described as having longer steady-state (V-ss), and the stress effect on spectral energy at higher frequencies was greater in the Argentinian vowels than in the Peruvian vowels. These results suggest that the F1 raising of the mid vowel in AS cannot be due to the undershoot triggered by temporal reduction (Lindblom 1963; Fourakis 1991) or by coarticulation effects (De Jong et al. 1993; Magen 1997; Lindblom et al. 2007), both of which are conditioned by temporal pressure or reduction. Thus, we interpreted this to mean that the mid vowel in AS

---

[61] Vowel space expansion can also be found variably cross-linguistically, possibly constrained by the inventory size of a language (Jongman et al. 1989; Bradlow 1995). For example, Bradlow (1995) found that the acoustic vowel space is larger in English (more phonemic vowels) than in Spanish (less vowels). Based on her results, Bradlow argues that, "a language-specific, base-of-articulation property" can play a role in determining the location of vowel categories. In addition to inventory size effects, a language-specific vowel space may also be affected by the distribution of allophonic realizations of phonemic vowels in a given language (Keating & Huffman 1984). It is an open question whether these factors (other than stress) are also involved in our data.

reflects hyper-articulation under stress (Lindblom 1990; De Jong 1995): possibly tongue dorsum raising for high vowels and jaw opening for low vowels. The variable strategies depending on vowel height are often interpreted to mean that paradigmatic contrasts are enhanced under stress effects. If this is the case for the mid vowel behavior in AS, the hyperarticulation hypothesis suggests that the mid vowel enhances a non-low feature (rather than non-high feature) and thus raises F1 under stress. Whether the dialectal difference is accounted for by the difference in CV-coarticulation effects or the hyperarticulation strategy, it is clear that the stress sensitivity of the vowel is different between the two dialects. Furthermore, this dialectal difference indicates that stress effects on the vowel are relatively greater in Argentinian Spanish than Peruvian Spanish, and this spectral difference is consistent with the asymmetric temporal behaviors in the two dialects (as shown in Figure 5.1)

To summarize, we have seen that acoustic properties of syllables are influenced by stress in Spanish, and the effect is significant in the realization of onsets as well as vowel. In addition, the results suggest that a syntagmatic relation between the syllable constituents (onset and rhyme) may be involved in the phonetic manifestation of stress.

5.2 Comparison of stress/accent and word position effects

A syllable can be perceived as prominent by virtue of any of several phonetic characteristics, such as greater loudness, greater duration, and greater pitch excursion. For a syllable in a prosodically strong position, prosody interactions have to ensure that relative prosodic strength can be expressed. In this section, we discuss how prosodic factors accommodate this prominence requirement through interactions. In Section 5.2.1, we compare the results of the stress effect to the accent effect, after which we can compare these prominence effects to position in word effects.

5.2.1 Prominence effects: stress and accent

The literature has been unclear on how different stress and accent manifestation is from other prosodic effects such as boundary effects. For instance, stress/accent effects can be marked by greater energy (Fry 1955, 1958; Lehiste 1970; Beckman 1986), and so can domain-initial effects (Cho et al. 2007). They can also trigger similar articulatory behaviors, and are categorized as the same kind of prosodic effect (e.g., extreme velum positions: Vaissière 1988). Differences have been observed as well. For instance, accent effects on vowels are more associated with articulatory expansion (Cho 2005), but this is not the case for domain-initial strengthening, which is argued as a local effect (Fougeron 2001).[62] In addition, Cho & Keating (2007) argued that boundary-initial effects may not be a local effect because vowels (intensity) were also affected by initial positions. Our results support, in part, the findings of Cho & Keating (2007) for English in that stress and accent effects are similar but still different from boundary initial effects. However, interaction patterns between the two types of prosodic effects in Spanish are different from what has been reported for English, suggesting that prosodic interaction patterns can be language-specific. We will discuss typological implications from our Spanish results which are consistent with but different in several ways from the English results of Cho & Keating (2007).

First, our results showed that stress has effects in duration that are independent of accent effects. The scope of the stress effect was the syllable, including onset. The results for /s/V syllables are shown in Figure 5.2, and can be compared to /x/V results in Figure 5.3.

---

[62] Cho (2005) examined both domain-initial strengthening and accent effects measured by tongue position. The result was also supported by his kinematic examination of lip gestures, which indicated that the accent effect is always associated with larger, longer, and faster lip opening and closing gestures, but this was not necessarily so for domain-initial effects (Cho 2006).

Figure 5.2 Stress and accent effects on C and V in the temporal dimension: /s/V syllables

(a) Syllable duration   (b) onset duration   (c) vowel duration
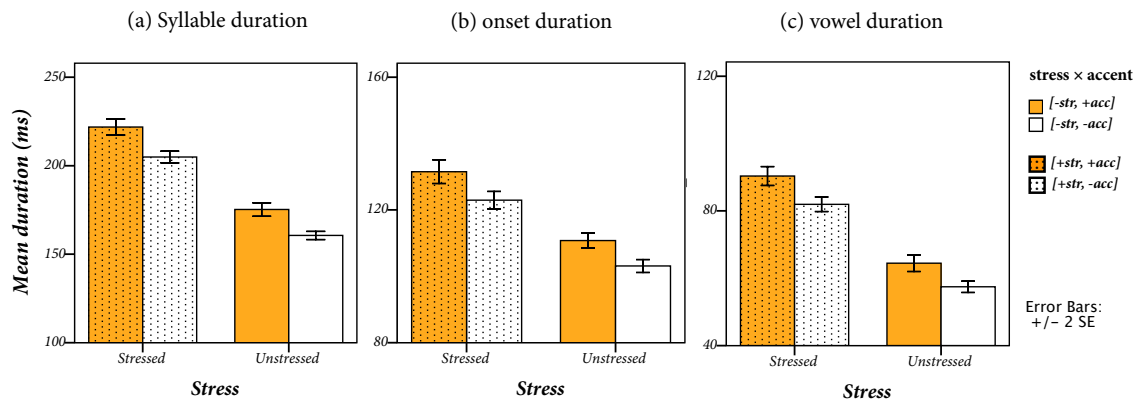
In these figures, the stress effect is clearly independent of the accent effect. The stressed onset and the vowel are lengthened when accented, suggesting that the accent effect can be independent of stress condition. Thus, both stress and accent influence not only the vowel duration but also the onset duration. The accent effect on the vowel is consistent in the results for /x/V syllables, but it is absent in the onset duration of /x/, as shown in Figure 5.3.

Figure 5.3 Stress and accent effects on C and V: /x/V syllables pooled across PoW conditions

(a) Syllable duration   (b) onset duration   (c) vowel duration
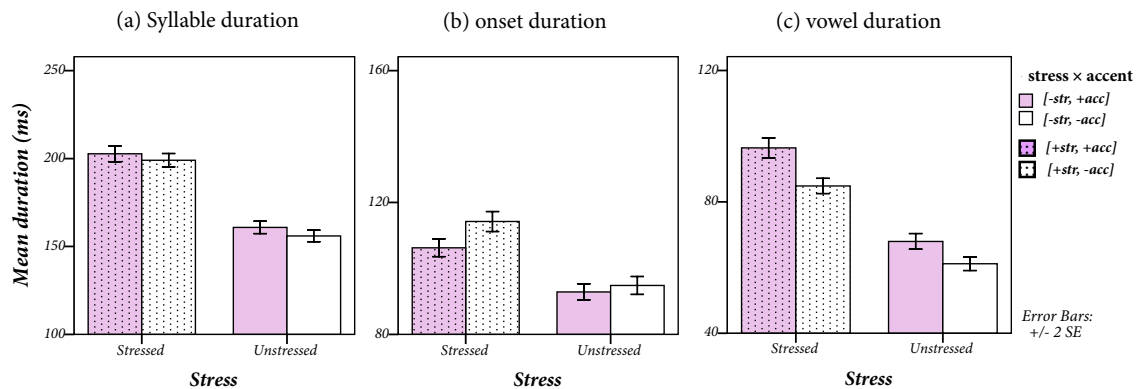
The insensitivity to accent in syllable duration (a) is due to the onset behavior in (b). That is, the duration of /x/ does not increase in accented position. On the other hand, the vowel durations in (c) clearly reflect the accent effect: e.g., stressed vowels are longer when accented. Thus, our

134

interpretation is that the accent effect can influence both onset and vowel, but the results can vary depending on the type of onset.

When stress and accent effects are compared for RMS intensity, we find that the accent effect is more consistently found in the vowel than the onset. The RMS results are summarized in Figure 5.4, separating the results by onset type.

Figure 5.4 Accent and stress effects in the intensity dimension: /s/V and /x/V syllables
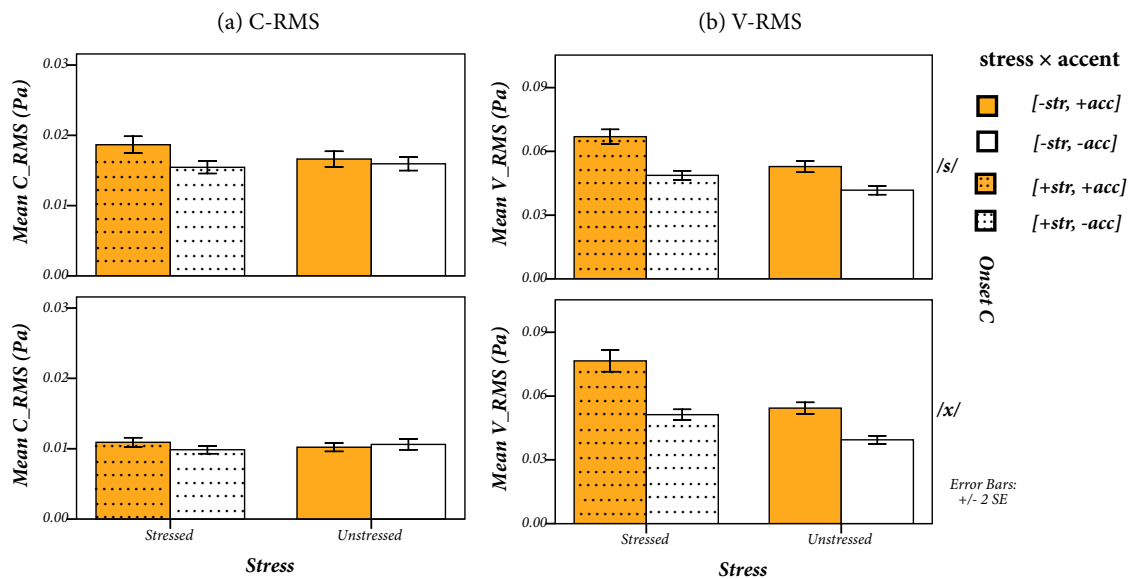
Figure (a) shows that accent mainly affects the stressed onset. Though the magnitude of the effect is smaller for /x/, the increasing RMS is present consistently for /s/ (top left figure). The RMS intensity difference is characteristic of the stress effect, and the accent effect enhances the difference, as shown in (b). The accent effect on the vowel is consistent between /s/V and /x/V syllables. Thus, the RMS intensity result shows two different interpretations with respect to syllable constituents. For consonants (onset), stress does not significantly affect intensity, but accent does, conditionally. That is, accent increases onset intensity, and this accent effect is significant only in stressed conditions. For vowels (rhyme), stress and accent effects in the intensity dimension are consistent and in the same direction, e.g., increasing intensity under stress and/or accent. This means that intensity would increase as the vowel is stressed and more

when the stressed vowel is also accented. This interaction between stress and accent in intensity was also found for the English data in Cho & Keating (2007). They discussed a possible account of increasing intensity as a characteristic of stress/accent effects, relying on the fact that faster jaw opening in CV can correlate with greater intensity (McClean & Tasko 2002) and that accented syllables in English do show faster jaw opening (Cho 2006).

Finally, we examined how spectral energy characteristics are influenced by stress and accent. For the onset, we review the first moment (COG) from the spectral moment analysis and spectral balance in /a/ for the vowel. The /s/V results represent the spectral patterns as shown in Figure 5.5.

Figure 5.5 Accent and stress effects in the spectral dimension: onsets/vowels in /s/V sequences

(a) COG of onset /s/   (b) Spectral balance in vowels
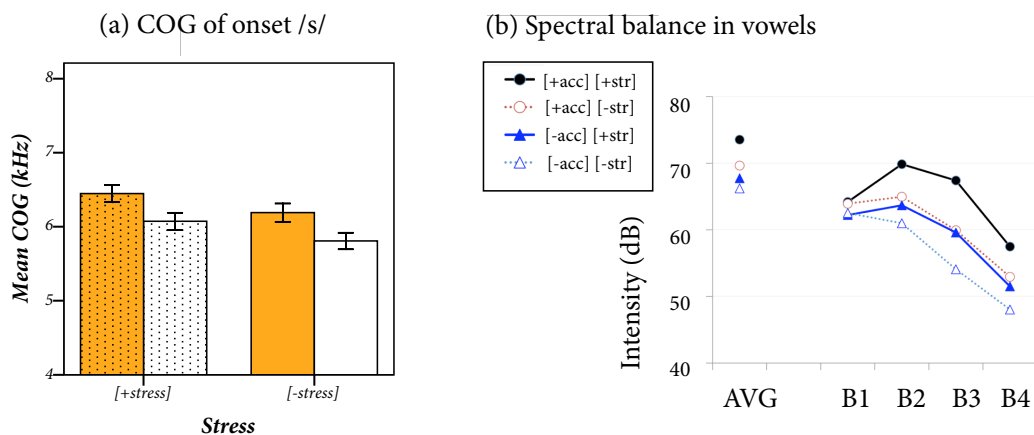
COG results are compared for combinations of stress and accent in (a). The stressed onsets show higher COG compared to the unstressed onsets, and both stressed and unstressed onsets increase COG consistently when they are from accented words. Note that the stress effect measured by COG is retained in both accented (dark bar) and unaccented (white bar) conditions. This means that the stress effect is well reflected in the onset fricatives when measured by COG. In addition, the accent effect can be cumulative, resulting in the highest COG values for the onset that is stressed and at the same time accented. All else being equal, decreasing energy at lower frequencies can also result in higher COG values. However, in our cases, the overall energy in the stressed and accented conditions increased, and thus the increase of COG can only be interpreted

as having more energy at higher frequencies. This cumulative nature of the prominence (stress and accent) effect can make the COG of an accented unstressed onset as high as that of the stressed but unaccented onset. Similarly, stress effects in the spectral energy distribution are differentiated within the accent conditions also for the vowels in (b). The spectral analyses of onset and vowel suggest that stress and accent effects can result in the same outcome in the spectral dimension. This also means that both stress and accent have spectral manifestations that, while independent, can be additive when onset and vowels are stressed and at the same time accented.

To summarize, in the temporal, intensity, and spectral dimensions, stress and accent affect onsets differently depending on the types of segments. Examination of stress and accent effects has revealed that the two effects can independently be manifested in multiple phonetic dimensions, and can enhance each other. Furthermore, it is evident that the complex nature of prosody can be understood in a compositional way when we disentangle the prosodic factors.

5.2.2 Word position effects in comparison to prominence effects

In this section, we will discuss how prominence (stress/accent) effects are similar to and/or different from position in word effects. Our results showed that the phonetic manifestation of word-initial boundary effects is indeed constrained by stress and/or accent conditions, and prosodic interactions among the factors ensure further that the isochronous rhythm can be retained.

We discuss first the temporal results of /s/V syllables (the results of /x/V are consistent unless otherwise noted). The duration of onsets is significantly longer word-initially, but the position in word effect disappears in the absence of stress and accent. This conditional effect is shown in Figure 5.6.

Figure 5.6 Position in word effects in relation to stress/accent effects: duration of onset /s/
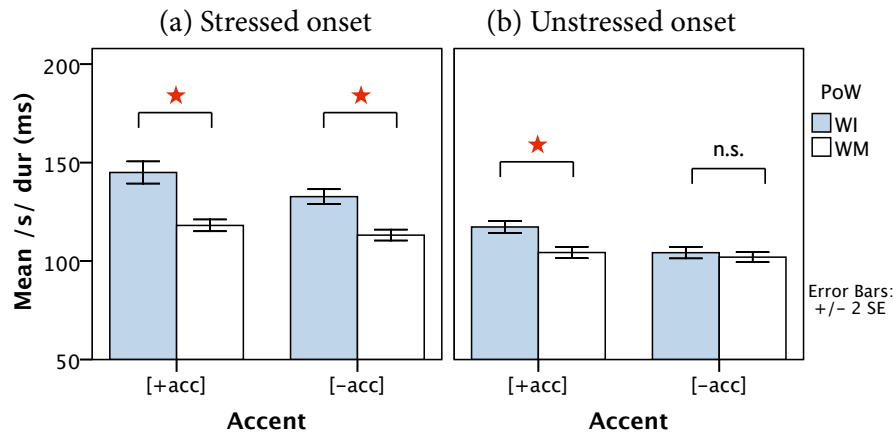
The duration of the initial /s/ is longer word-initially when accented. This initial lengthening does not occur when the onset is unstressed and at the same time unaccented, resulting in the boundary effect being present only in the stressed condition in (b) in the absence of accent. This means that either stress or accent has to be active in order to have the word-boundary effect. This conditional boundary effect on the prominence factors (stress/accent) was also found in English (Turk & Shattuck-Hufnagel 2000; Cho & Keating 2007), suggesting that boundary effects may be constrained by prominence factors in some phonetic dimensions. This raises an open question as to whether the conditional nature of the boundary effect can be language-specific or -universal. Future studies looking at prosodic interrelations will allow us to address more definite answers.

Our temporal results indicated a possible ceiling effect also discussed in Cho & Keating (2007); i.e., a restriction or limit on total strengthening of a single segment. This ceiling effect was found in Cho & Keating for domain-initial strengthening such that boundary strengthening was greater when other prominence effects were absent. When the domain initial strengthening factor was concomitant with other prosodic factors, then the initial strengthening was much less. In our results, syllable durations may be bound by a ceiling effect, which makes sense if we characterize the rhythmic nature of Spanish as having higher-level prosodic structure. We will demonstrate this phenomenon by comparing individual and combinatory effects. For the purpose of comparison, we use a prosodic index that represents the number of prosodic factors

involved (0~3) and the specific prosodic factor(s) activated (P=PoW, A=Accent, S=Stress) from "None" (0: no active prosodic factor) to "All" (3-PAS: all three factors are active).

We first discuss prosodic effects in the durations of C, V, and syllable, and we see how individual or combinatory prosodic effects are responsible for the lengthening of syllable duration.

Figure 5.7 Individual and combinatory prosodic effects: duration
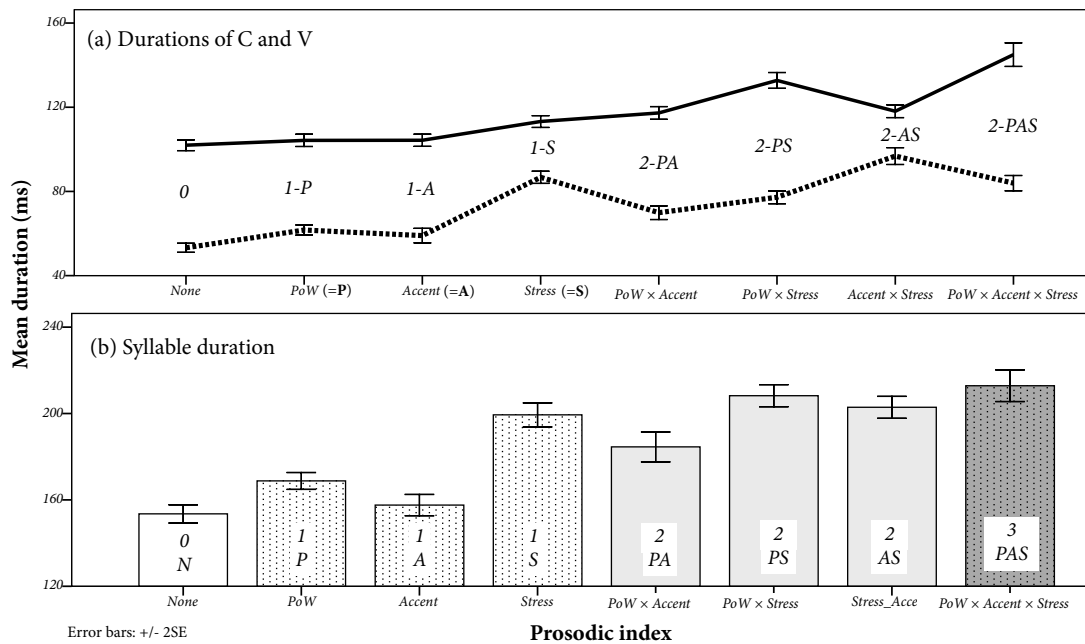
As described earlier, the prosodic index consists of the number of factors involved (0~3) and the active prosodic factor(s). We will limit our discussion to the comparisons that give us clear interpretations regarding prosodic interactions. Firstly, among the single factors (1-P, 1-A, 1-S) in the top figure, the stress effect (1-S) is the greatest in the durational lengthening of both C and V (Note that Stress (1-S) condition represents "Word-medially located unaccented stressed" syllables). On the other hand, accent (1-A) and position in word (1-P) mainly lengthen consonant duration, though these effects are relatively weaker compared to the stress effect (1-S). This comparison among the single factor effects enables us to interpret the syllable duration results in (b). For instance, the lengthening of PoW condition (1-P) in (b) is attributable mainly

to the consonant lengthening, whereas the stressed syllable (1-S) in (b) becomes longer since both C and V are lengthened. Secondly, stress effects show an interesting pattern. As a single prosodic factor (1S) in (a), it has the greatest lengthening effect on both onsets and vowels. When the stress effect is combined with other prosodic effects (2-PS, 2-AS, and 3PAS), the lengthening pattern of C and V are not consistent. Interestingly, the total durations of syllables from these interactions are alike: 2PS ≈ 2AS ≈ 3PAS) in (b). In particular, when accent and stress are concomitant (2AS), the vowel lengthening was relatively smaller but the consonant lengthening was relatively greater. This suggests that the durations of the syllable constituents compensate for each other within the syllable. The adjustments among prosodic effects are also clear in the PAS condition, where the vowel lengthening was greater than in the other two factor conditions, and in this case, the consonant lengthening was relatively weaker. This seems to indicate a possible ceiling effect on the syllable duration. What we can learn from these interactions is that prosodic effects do interact in a specific way (either conditional or additive), and that the interaction patterns are constrained by the total lengthening allowed for a syllable. We interpret this to mean that, in Spanish, total syllable durations affected by prosody may not be allowed to exceed a certain temporal reference (e.g., constrained by isochronous rhythm in Spanish).[63]

The second dimension we investigated was intensity; we will compare how individual or combinatory effects are active in the intensity manifestation of prosody. We used C/V ratio values as referents in comparing prosodic effects. The general pattern was that the C/V RMS ratios decreased when syllables were either stressed, accented, or word-initial syllable (see Figure 4.6). This pattern was consistent among the three prosodic effects and thus was interpreted to mean that prosodically strong syllables would show relatively greater intensity in the vowel. In the following figure, the C/V RMS ratio results are compared for the PoW-only context (1-P), PoW with either Stress (2-PS) or with Accent (PA), and PoW with both (3-PSA).

---

[63] In Cho & Keating (2007), the found a ceiling effect such that domain-initial strengthening is greater in the absence of stress/accent effects. This result is different from ours since the position in word effect is also absent when stress/accent is inactive.

Figure 5.8 PoW effects measured by C/V RMS ratio in combination with stress and/or accent
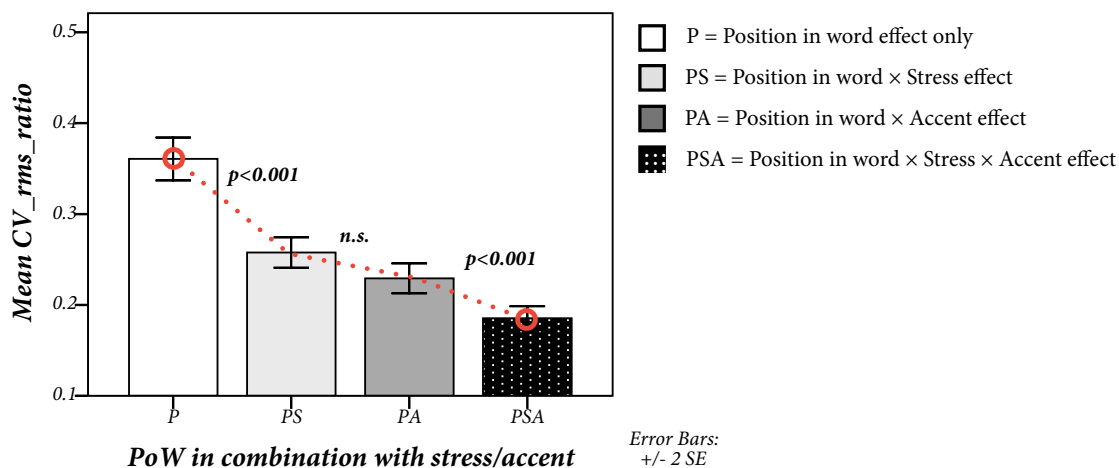
The C/V RMS ratio is the greatest when position in word is the only active factor (1-P), because PoW affects C-RMS as well as V-RMS in our data.[64] In comparison, the other prosodic effects were mainly in the vowel intensity. As more prosodic factors were combined, the ratios decreased, which we interpret as strengthening of syllables. Finally, we find two-step changes in C/V RMS ratio (P < PS, PA, AS < PSA) in the given figure. The similar effect found in the C/V ratios between AS and PS suggests that stress and accent effects are comparable when they are combined with the boundary effect. We can understand that both stress and accent are indeed a similar kind of prosodic effect in the intensity dimension, and also that the two types of effects (prominence vs. boundary strengthening) can be additive as we can see in the decreasing RMS ratios from P to PSA in the figure.

The result of C/V ratio in RMS reveals an interesting parallel between sonority and prominence manifestation in intensity. Lavoie (2001) discussed how C/V ratios in RMS intensity correlate with the sonority scale. She found that RMS ratios between C and V show a trend by manner of articulation, and the patterns match approximately with positions in the sonority

---

[64] The fact that PoW can affect both C and V was also found in Cho & Keating's (2007) results in English. They found that vowel intensity was affected by domain-initial conditions, which led them to conclude that word-initial strengthening may not be strictly local.

hierarchy (e.g., obstruents > nasals; fricatives > approximants). This can be interpreted to mean that a greater contrast in intensity between C and V may correlate with the general preference for low profile sonority for onsets. In parallel, we found that the C/V ratio changes in intensity may represent prosodic strength of syllables. That is, individual changes of C and V in intensity may not directly signal prosodic strength. Instead, the changes in C/V ratios in intensity seem to correlate with prosodic strength such that a greater contrast in intensity between C and V is found when a syllable is associated with more positive prosodic factors (stress×accent vs. stress×accent×PoW). It is possible that two syllables can have similar C/V ratios either by having increased intensity of vowel with the intensity of an onset being intact, or by increasing intensity of both C and V with a greater intensity of the vowel. In the first case, we may not see a contribution of the onset intensity to the syllable prominence since the syllable prominence can be expressed by the intensity contrast between C and V. In the second case, on the other hand, we see intensity increases in both C and V. The following figures illustrate the role of intensity expressed by C/V ratio in describing the parallel relationship between a schematized sonority scale and the strength of prosodic effects. A schematized sonority scale in Figure (a) is compared to prosodic strength in Figure (b), as expressed by C/V ratios in intensity. In these figures, variable C/V ratio values in intensity are represented by "Δa, Δb, Δc".

Figure 5.9 C/V ratios in RMS in relation to schematized sonority scale and prosodic strength

(a) Sonority scale & RMS C/V ratio
(Among different manners of articulation)
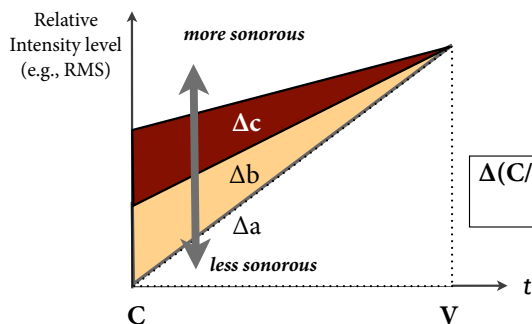
(b) Prosodic strength & RMS C/V ratio
(Within the manner of articulation: fricatives = /s/ or /x/)
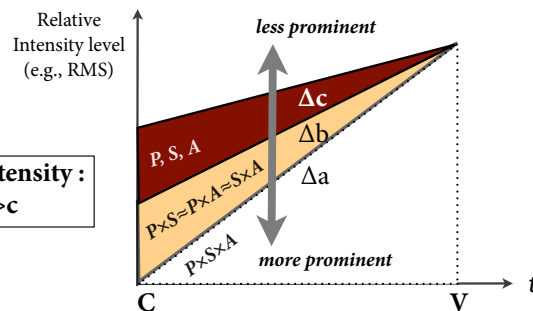
The intensity tilt (Δa, Δb, or Δc) in these figures describes the intensity contrast between C and V. In Figure (a), the C/V ratios in intensity ($\Delta a > \Delta b > \Delta c$) describe a schematized sonority scale among different manner of articulations of onsets. As discussed in Lavoie (2001), the sonority scale tends to correlate with C/V ration in intensity (RMS). In the given figure, 'Δa' can characterize syllables with a low sonority profile of onsets (e.g., stops), while 'Δc' characterizes syllables with a high sonority profile of onsets (e.g., approximants). Thus, a greater contrast in intensity between C and V (e.g, Δa) is the outcome of having an onset of a low sonority profile while a smaller contrast between C and V (e.g., Δc) is the syllable with an onset of a high sonority profile. This relation holds also for prosodic strength described in Figure (b), where prosodically prominent syllables are linked to the smaller value of C/V ratio in intensity. This connection may explain why a low sonority profile is preferred for onsets in general: that is, increasing contrast in intensity between C and V can be made by selecting a low sonority onset, and the syllable can be interpreted as being more prominent.

Alternatively, C/V ratios in intensity can vary by changing vowel intensity while the intensity of onsets is kept constant. This is illustrated in Figure 5.10.

Figure 5.10 Constant and variable C/V ratios in RMS and prosodic interpretations.

(a) Variable prosodic strength expressed by variable RMS C/V ratios
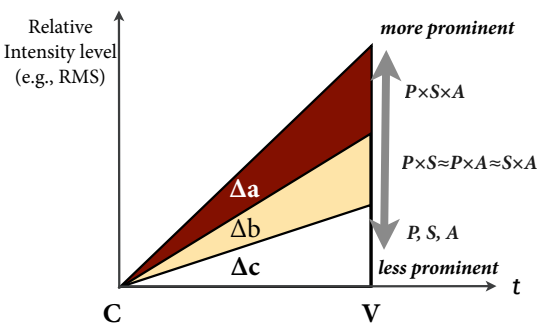
(b) Same prosodic strength expressed by same RMS C/V ratios
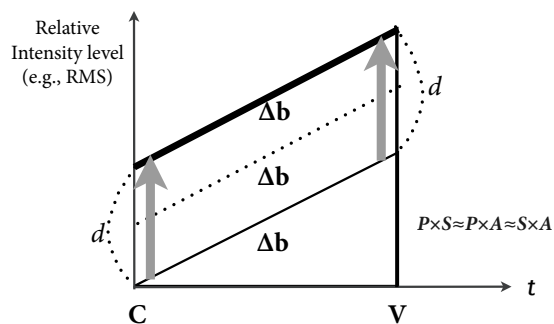
Figure (a) describes how variable C/V ratios in intensity correlates with the prosodic index. That is, the more prosodic features are active, the greater contrast in intensity is made between C and V. For example, PSA corresponds to 'Δa', which is greater than Δb or Δc. In this Figure (a),

intensity changes are made only in V, but the ratio also changes accordingly. This illustrates a situation where vowel intensity increases affected by prosody without making intensity changes in onsets. Figure (b) shows another possible intensity manifestation within syllables. In this case, both C and V increase intensity but proportionally between the two. Regardless of the increasing intensity by '*d*' in Figure (b), the three lines represent identical C/V ratios in intensity, and they represent the same category, here 'PS, PA, or SA'. The lesser the onset sonority is, the greater the intensity contrast between C and V can be made, all else being equal. What this means is that syllables can achieve greater prominence by selecting an onset with a lower sonority profile and without changing the quality of a vowel. This description holds also for prosodic strength in (b) grouped by the prosodic index (one-way, two-way, and three-way prosodic combinations). In this figure, a greater change in C/V intensity ratios represents a stronger prosodic effect in the intensity dimension: Δa (PSA)< Δb (PS, PA, SA) < Δc (P, S, A). This suggests that the intensity contrast between C and V can express how syllables are affected by prosody in terms of intensity. Possible roles of intensity contrast between the syllable components deserve more attentions in the future particularly with the question of how the ratio changes in intensity contribute to the perception of syllable prominence.

The third dimension we discuss is spectral properties. We compare individual and combinatory prosodic effects by COG changes in fricative onsets (/s/), and by spectral balance in vowels. The main generalization regarding strong prosodic effects measured by COG was that COG values increased when prosodic factors were active. For the spectral interpretation, increasing energy at higher bands rather than overall intensity signals prosodic strength. Thus, we will interpret increasing COG and increasing higher band energy as strengthening effects of prosody on the onset and the vowel, respectively. The following figure shows the changes in COG of onset /s/ by individual prosodic effects and combinatory effects using prosodic index ([0~3]×[S, A, P]), as used earlier. The results shown in the following figure are pooled across dialects.

Figure 5.11 COG comparison by prosodic index: onset /s/
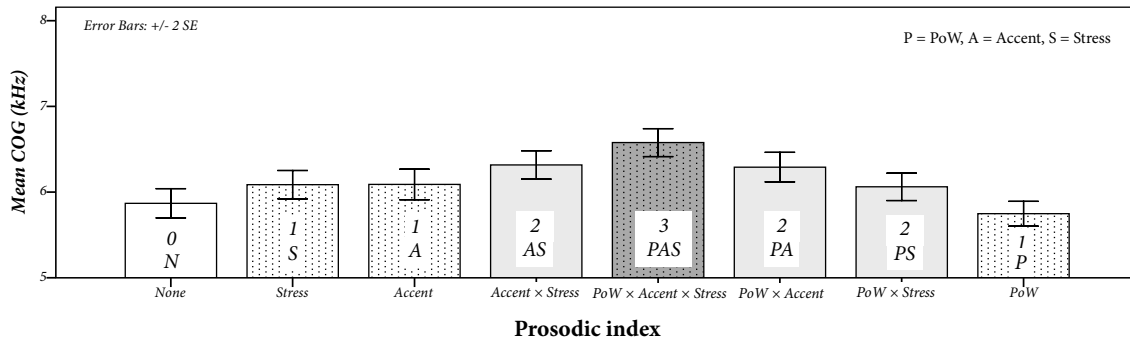
In this figure, PoW effects (1-P) are absent in changing COG of the fricative. As a single factor, stress (1-S) and accent (1-A) effects are similar. These two factors increase COG when they are combined (2-AS). Although the word position does not have an effect on its own, it increases COG when combined with stress and/or accent. This suggests that PoW is conditioned by stress or accent, and the combinatory effect is greater with accent. Thus, the COG value is the greatest when the three factors are combined (3PAS); it is weaker when only two of the factors (2AS, 2PA, 2PS) are combined; and it is the weakest when the position in word factor is not incorporated with the prominence factors (identical to 0 condition). Thus, we can conclude that prominence effects change COG of onsets while the initial-boundary factor has an effect only when a prominence factor is activated.

Stress and accent as prominence factors have similar effects on spectral energy distribution of vowels. In the following figure, the averaged spectral balance of /a/ is compared among prominence-only (1S, 1A, 2AS), boundary-only (1P), combination of the two (2PA, 2PS), and combination of all (3PAS) contexts. Intensity is measured at five frequency bands: B1 (0~0.5 kHz), B2 (0.5~1kHz), B3 (1-2 kHz), B4 (2-4 kHz), and B5 (4-8 kHz)

Figure 5.12 Prominence and boundary effects comparison: spectral balance of /a/
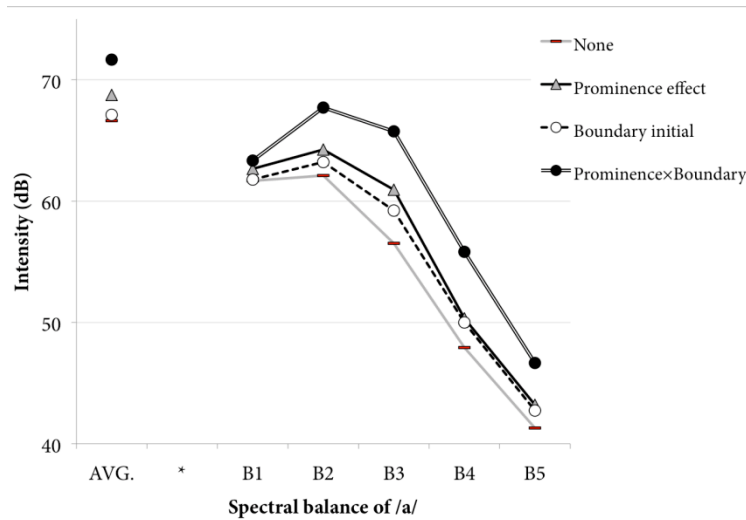
The overall energy distribution from B2~B5 can differentiate the three contexts (none, prominence-effect-only, all), but not between boundary- and prominence-only contexts. The boundary initial effect tends to increase spectral energy at higher frequency bands, but not as significantly as in the prominence-only condition. This suggests that boundary and prominence effects are distinct. When all three factors are combined, the effect was the greatest and significantly different from the other contexts. Thus, we conclude that the three prosodic factors enhance each other, even though the word boundary effect seemed to be weak on its own.

Understanding interrelations among prosodic factors and possible compositionality of prosodic effects can provide us insight into the grammatical structure of prosody, which is often reflected in listeners' perceptual sensitivity to variable phonetic cues (Dupoux et al. 1997; van der Hulst 1999).[65] We also know that listeners can make use of different prosodic cues in speech segmentation (e.g., Tyler & Cutler 2009). This indicates that listeners are tuned to fine-grained phonetic details based on their knowledge of prosody manifestation at different levels. Tyler & Cutler (2009), for example, investigated how listeners' of Dutch, English, and French make use of

---

[65] The underlying mechanism of the phonetically varying patterns of prosodic manifestations leads us to an open question: what information a language community has to share in order for them to correctly produce/perceive linguistic signals for stress, whose interpretation is often hard to master for non-native speakers.

suprasegmental cues in segmenting sound sequences. They found language-specific sensitivity to pitch variation depending on position, such that English listeners attended to pitch variation occurring in left-edge position whereas French listeners used pitch cues only in right-edge position. For the Dutch listeners, both positions were informative since pitch variation in Dutch is not restricted by position. These results suggest that interrelations among prosodic factors also have to be part of the native speakers' knowledge. Thus, we can learn from prosody interactions about how prosodic structure can influence perceptual behaviors. Further studies on this topic should be given more attention in the future.

To summarize, we have discussed how stress, accent, and word position effects can be compared in their phonetic manifestations. Both stress and accent consistently lengthen the vowel and the two effects can be additive (up to a certain limit). On the other hand, onset lengthening seems to be ensured by one of the two prominence factors (either stress or accent). When the two factors co-occur, vowel lengthening is greater than the effect of the individual factors, but this is not the case for the onset lengthening. That is, no additional lengthening occurs in the onset even if stress and accent are active together. Interestingly, the syllable durations become similar between stressed (but unaccented) and accented (but unstressed) syllables. This implies that lengthening effects are ensured either by stress or by accent, but there can be a ceiling effect in manifesting syllable durations possibly constrained by a higher-level rhythmic property in this language.

5.3 Segmental/Subsegmental variation due to stress and rhythmic effects

So far, we have found some evidence to support the idea that isochronous rhythm as higher-level prosody constrains the manifestation of the word-level (stress) prosody. This is shown in the asymmetric syllable interval variation between the two dialects, a ceiling effect in the total durations of syllables, and interaction patterns among prosodic factors. In Section 5.3.1, we discuss the isochrony characteristic of Spanish as it relates to stress manifestation. This is based on the idea that stressed syllables may need to comply with the so-called syllable-timedness rhythm as has been often mentioned to characterize Spanish differently from other stress-timed

147

languages including English. In Section 5.3.2, we provide a gestural account of the syllable internal temporal variation found in the two dialects.

5.3.1 Prosodic structure, prosodic interaction, and isochrony

While no complete set of linguistic criteria has successfully established the isochrony dichotomy in the literature (e.g., Roach 1982; Dauer 1983; Bertinetto 1989), a scalar degree of isochronous rhythm can characterize one group of languages from another (Dauer 1983; Bertinetto 1989). On this typological continuum, we assume that Spanish is closer to the 'syllable-timed' end, rather than the 'stress-timed' end. What, then, would happen when two dialects of a language that is characterized as syllable-timed show (sub-) segmental variation, which is not simply attributable to dialectal differences. Argentinian Spanish was investigated since it has been impressionistically described as having a more "rhythmic" pattern than other dialects of Spanish (Colantoni & Gurlekian 2004). Varieties of one language can differ in rhythmic structures as much as two different languages. We do find dialectal variation in rhythmic regularities in the literature. For instance, Major (1985) claimed that formal Portuguese shows properties that are typical of syllable-timed languages and informal Portuguese shows different rhythmic characteristics that could be close to stress-timed languages. In addition, some southern varieties of Italian are described as having different rhythm from standard Italian, which is frequently taken as one of the 'syllable-timed' languages (Grice et al. 2005). Differences in rhythm across varieties of English have also been reported (e.g., Singapore English compared to British English; Low et al. 2001). Taken together, prosodic variation within languages can also be substantial.

We have seen that the two dialects show asymmetric stress effects between syllable constituents when measured by the narrow method, suggesting that micro-variation within a syllable is also possible. We found an asymmetric pattern of stress effects between the two dialects such that sub-phonemic components were lengthened asymmetrically: longer C-ss for Peruvian Spanish vs. longer V-ss for Argentinian Spanish when syllables are stressed. The durational difference in our findings suggests that Argentinian Spanish could differ phonetically in the intra-segmental modification within the syllable, which may explain why Argentinian Spanish

sounds rhythmically different even though the durational modifications (C and V) by stress are similar. We find evidence to support this interpretation in the literature. For example, White et al. (2009) showed that multiple factors including fortition/lenition of segments (consonants and vowels) contribute to the perception of rhythmic differences. This means that rhythmic characteristics are more than durational regularities, and can be reflected in the phonetic realizations of segments. Thus, the asymmetric pattern found within the syllable in our data may be understood by referring to isochronous rhythm shared by the two dialects. That is, the two dialects are similar in that both preserve isochronous rhythm at high level of prosody, and differ from each other in that the vowel contribution is relatively greater in Argentinian Spanish, whereas the consonant contribution is relatively greater in Peruvian Spanish.[66]

5.3.2 A gestural account of the temporal manifestation of prosody

The dialectal comparison in the temporal results of prosodic manifestation asks for an account of the within-syllable temporal modifications that characterized the two dialects of Spanish. The following section provides a gestural account of this asymmetric pattern. The results on temporal modulation revealed an interesting intra-syllable difference between the two dialects. In Peruvian Spanish, the steady-state of onset fricative (C-ss) tends to be longer and the vowel steady-state (V-ss) to be shorter in stressed CV syllables, compared to Argentinian Spanish. This syllable-internal asymmetric pattern calls for an explanation of how intra-syllabic modification can be modeled. We will strengthen gestural coordination within the syllable employing a prosodic gesture, the "pi-gesture", to develop a possible account.

Gesturally, consonant and vowel are in-phased (Browman & Goldstein 1989, 2000). This means that C and V are coordinated in time, and initiated together in gestural planning. The following diagram roughly represents this gestural coordination between C and V within the syllable (Nam 2004).

---

[66] One intriguing question is what makes one dialect choose one direction and not the other. This is an open question for future study.

Figure 5.13 A schematized phasing relation within the syllable
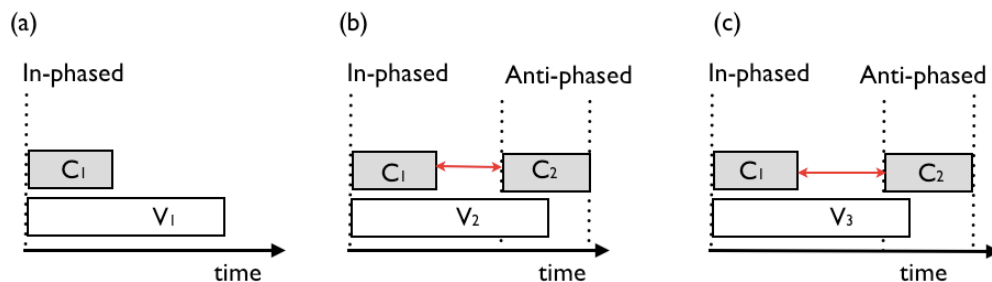
For example, the onset consonant (C_1) is in-phased with the vowel, as shown in (a)~(c). On the other hand, the coda consonant (C_2) is anti-phased with the vowel, or sequential in time, as in (b) and (c). In this configuration, the onset C is less likely to influence the duration of V than coda C may, given the gesural time (length of the gestural score) and magnitude (height of the gestural score). That is, the onset C cannot change the vowel duration unless the C gesture modifies its timing specification. On the other hand, the coda consonant (C_2) in (b) is phased sequentially over V, and the vowel duration in the output can change depending on where the coda C is phased: compare (b) and (c) to see the vowel durations that differ as a consequence of the C_2 phasing within the syllable. The arrows in (b) and (c) indicate the potential vowel duration. This gestural phasing relationship nicely predicts a potential compensatory relationship between vowel and coda consonant, as discussed in Nam (2004). However, this account does not make a prediction on its own regarding onset and vowel variability.

A possible model for this pattern can be given by adopting a prosodic gesture. The prosodic parameter adopted in this model is the "pi(π)-gesture", which was introduced to model prosodic boundary adjacent lengthening phenomena (Byrd & Saltzman 2003). Pi-gesture modulation as used there slows the gestural timing at or near a prosodic boundary and the phonetic output of this modification can be, for example, phrase-final lengthening. Kim & Nam (2009, 2010) proposed that a prosodic gesture could play a role also at the syllable boundary to account for within-syllable variation. We will review the proposal and discuss how it can account for the differences of sub-phonemic lengthening between the two dialects.

The Task Dynamics Application (TADA), developed at Haskins Laboratories (Nam et al. 2004), allows us to simulate output speech patterns based on gestural specifications as input. The application is structured with gestural scores, each of which is specified for sequences of speech gestures in time, magnitude, and phasing. In addition, the application implements prosodic gestures as an independent parameter, and therefore we can combine segmental gestures and prosodic gestures within a simulation. This computational simulation is summarized as a diagram in Figure 5.14, where the output is the gestural score after the application of the pi-gesture.

Figure 5.14 Simulation in TADA (Task Dynamics Application)
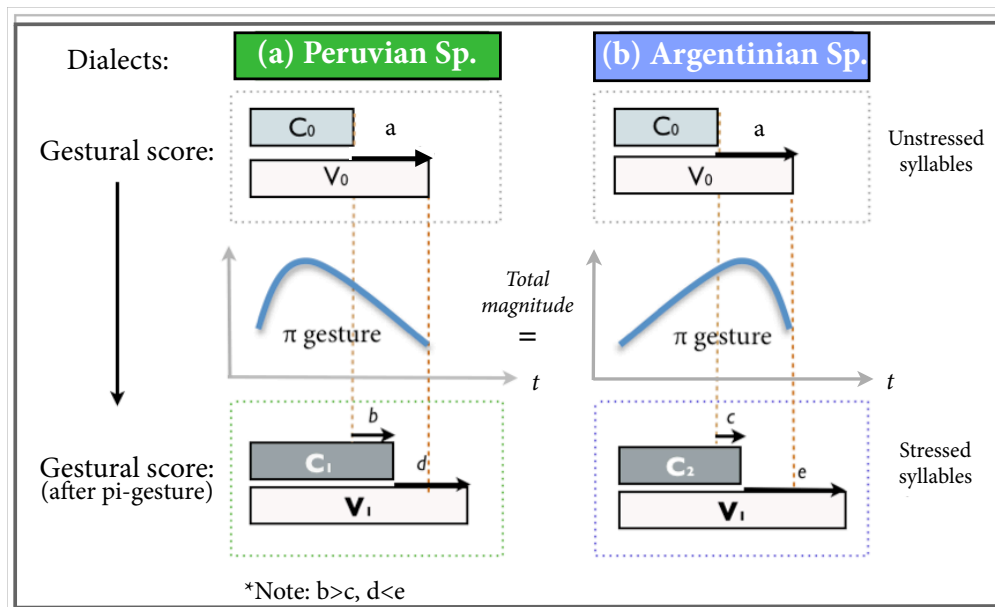
Adopting the proposal originally made for prosodic gestures, or pi-gesture (Byrd & Saltzman 2003), a modification of the gesture in time and magnitude within the dynamic system TADA was implemented in this simulation to capture the timing patterns that are found between the two dialects. This modification is represented as a skewness of the pi-gesture in the middle of the figure.[67] Note in the figure that the gestural specification for unstressed syllables is identical for the two dialects, and also that the total magnitudes of pi-gesture is identical in the two

---

[67] Pi-gesture parameter is otherwise specified with a gestural score box that appears also in this figure to represent C and V gestural scores. This modeling is based on Kim & Nam (2009, 2010).

specifications. The output of the specification is two types of stressed syllables as shown at the bottom of the figure, which differ in the durations of C (C1>C2) and V (d<e) on the surface. This simulated model correctly distinguishes the two patterns that can characterize the two dialects. For example, the left-skewed pi-gesture predicts positive onset strengthening whereas the right-skewed pi-gesture for vowel strengthening. The first pattern represents the stress manifestation in Peruvian Spanish and the second pattern for Argentinian Spanish.

The next step would be to gather articulatory data to compare to model simulations. In addition, several important questions remain for future research. First, the asymmetric prosody manifestation found in this study opens the question of why one dialect takes one pattern (e.g., left-skewed prosodic gesture in PS) and not the other (e.g., right-skewed prosodic gesture in AS). Language internal (e.g., prosodic interactions) and external factors (e.g., dialectal contact) need to be considered to answer this question. The second question concerns modeling prosody. We want to know whether there are other factors that can influence or determine prosodic gesture patterns, and what could be the linguistic consequences of variation in these prosodic gestures. The prosodic model discussed in this study can provide typological predictions, which can be tested through empirical investigation in the future.

Chapter 6 Conclusion

This dissertation investigated [1] how lexical stress is phonetically manifested within the syllable, [2] how stress manifestation interacts with other prosodic effects, such as phrasal accent and word-boundary effects, and [3] how rhythmic characteristics of a language are involved in prosodic interactions.

We have found that stress in Spanish has its own phonetic manifestation independently of other prosodic effects, particularly in the temporal and spectral dimensions. Among the three prosodic factors, the stress effect as a main factor was the most consistent and the greatest in the phonetic dimensions investigated. The results also showed that the three prosodic factors (lexical stress, phrasal accent, and word-boundary) interact in the duration, intensity, and spectral dimensions. We observed finally that higher-level prosody beyond the phrasal level possibly constrains lower-level prosody manifestation. If this hierarchical relation is in effect, then variability of segments at any given lower level of prosody can also be constrained by the prosodic requirement at higher levels. That is, stress effects may not be guaranteed to be greater than other effects when accent/or boundary effects are concomitant, since all, as lower level prosodic components, they all have to ensure that the upper bound constraint is respected. This can explain why there is no linear relationship among prosodic effects, and instead, we find complicated interrelations among prosodic factors that do not seem to be language-universal. In this dissertation, we considered higher level prosody (higher than phrasal level) as an important factor to account for apparent language-specific prosodic interactions. Further investigations will shed additional light on how multiple prosodic factors are interconnected within the prosodic system of a language, and how prosodic components accommodate prosodic constraints. The disentanglement of prosodic effects leads us to a better understanding of the complex nature of prosody.

The results on prosodic effects were described in this dissertation provide several linguistic implications that are important for future research. Firstly, syllable prominence may be expressed by relative phonetic values (such as C/V ratio in duration and intensity) between

153

syllable constituents. For example, the prosodic effect within the syllable could be well characterized in the intensity ratio analysis: e.g., smaller C/V ratio correlates with prosodic strength. While the phonetic effect within the syllable varied a lot such that individual or combinatory effects do not pattern consistently, decreases in both C/V durational and RMS intensity ratios correctly describe positive prosodic effects in a cumulative way. This emphasizes the fact that the prominence of syllables is interpreted not by the absolute amount of phonetic changes, but by the relative strength between syllable constituents. In other words, prominence can be better understood in terms of relative strength between syllable components. One implication here is that while we did observe some complimentarity in stress effects on onset and vowel, it is of course also possible for greater duration and/or energy in the vowel to make syllables prominent even when the phonetic properties of onsets stay intact. This interpretation needs further investigations using perceptual data. More research along these lines will help us to build up a complete picture of relational strength of prosody.

Secondly, the word boundary effect is different from prominence effects (stress and accent), and it may even be conditional to the prominence factors. We have seen that stress and accent effects, grouped as prominence effects, showed a distinct behavior from the word-level boundary initial effect. The overall prominence effect tends to affect the vowel in a more consistent way than it does for the onset. The narrow measurement method provided insight into how sub-segmental units can be affected by prosody, and can be linguistically encoded in the manifestation of prosody. We also found that the word boundary effect is often conditioned by prominence factors. That is, the word boundary effect is not strong in the absence of prominence effects, but becomes a strengthening effect combined with prominence effects. This result is different from other study (e.g., Cho & Keating 2007 for English) in that the boundary effect is the greatest when prominence effects are absent. Cross-linguistic data in the future will allow us to determine whether conditioning nature of the boundary effect may be language universal, and to determine how languages differ in conditioning boundary effects in relation to prominence effects.

Thirdly, temporal modulations within a prosodic domain (i.e., syllable) can vary even at the level of dialects. Here, the asymmetric contribution between syllable constituents was found

in the temporal manifestation of stress, and the possible asymmetric lengthening was accounted for based on a gestural representation. The apparent complementarity between consonant and vowel can be attributed to two structural motivations. On the one hand, syllable constituency is involved in the variation, and on the other hand, higher-level prosody (e.g., isochrony) can play a role as an upper bound in constraining variability. The gestural account discussed in this dissertation demonstrated how stress manifestation could vary within the syllable. Further empirical data will help us to develop a complete prosody model, through which typological variability in prosody manifestation can be examined. In addition, establishing the phonetic and phonological connections between sub-segmental variation and prosody would require similarly detailed phonetic examinations into the interactions among language- and/or dialect-specific prosodic interactions/interrelations.

The final implication is related to the question of why prosodic effects are rarely phonologized. Given the fact that stress can influence the phonetic realization of both onset and vowel, we may ask why the possible variation of onset is not reflected in the phonology in general. The same question is also addressed in Lavoie (2001: 167): "How do such robust effects (stress effects) escape phonologization (given the consistent phonetic alternations of consonants influenced by stress)?" Her explanation is that word initial position is phonologically more important and the strengthening in that position "outweighs" weakening in medial position. This can be further interpreted to mean that that phonology in the grammar refers mainly to the segmental phenomena occurring in prosodically strong position (e.g., stressed and/or word-initial position). The results in our current study provide an alternative way to view the rarity of onset roles in the grammar. Prosodic effects are gradient but not always cumulative; the effects partly depend on which prosodic factors are involved and how the multiple prosodic factors interact in a specific phonetic dimension. For example, onset lengthening does not necessarily occur when consonants are word-initial. We have seen that the word-initial lengthening of onset is conditioned by prominence factors. In addition, vowels following word-initial onset are also affected in duration and intensity. As we have seen in the ratio analysis of intensity, the word-initial prominence can be achieved by having a greater contrast in intensity between C and V even without affecting onset intensity. Thus, the interrelations among prosodic factors are one of

155

the key questions in answering what the role of prosody in the grammar can be. Although prosodic effects are indeed rarely phonologized, we find evidence of an important role of prosody in sound change. One example is the irregular sound change that was once considered as an exceptional pattern in the First Germanic Sound Shift (Grimm's law). The exceptional case was explained by Verner's law (Verner 1877): the voiced stop occurrences in the Proto-Germanic (Proto-Germanic) language happened in the onset of unstressed syllables, which were fed by accent shift to word-initial position in PGmc. Without making reference to prosody, the unified historical phonological process could not have been recognized and would have been left as a mystery. The research questions taken up in this dissertation lead us to a better understanding of the role of prosody as one of the fundamental conditioning factors in segmental variation.

In conclusion, this dissertation showed how rhythmic characteristics in a language may constrain prosodic manifestation at lower levels. Based on the results, we understand that prosodic factors interact in a way such that the outcome of the interaction can still be faithful to higher-level constraints. These findings indicate that stress manifestation as well as rhythmic typology (e.g., syllable-timedness) is unlikely to be categorical, as other prosodic factors also play a role in the realization of segments on multiple phonetic dimensions. Further research can address how native speakers become attuned to such fine-grained prosodic manifestations, and to what degree listeners can take advantage of such phonetic properties. It is hoped that this dissertation contributes to a better understanding of prosody, as a well-structured grammatical component, and its connection to various segmental phenomena in languages.

References

Abaurre, Maria B. M. & Flaviane R. Fernandes-Svartman. 2002. Secondary stress, vowel reduction and rhythmic implementation in Brazilian Portuguese. In *Fonologia e Variação: Recortes do português brasileiro*, eds. Bisol, L. and C. CBrescancini, 54-81. Porto Alegre: Editora Universitária da Pntifícia Universidade Católica do Rio Grande do Sul.

Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Amador-Hernandez, Mariscela. 1986. Spanish as a "syllable-timed" language. *Journal of the Acoustical Society of America* 80: S96-S96.

Astruc, Lluïsa & Pilar Prieto. 2006. Stress and accent: Acoustic correlates of metrical prominence in Catalan. Paper presented at *ISCA Tutorial and Research Workshop on Experimental Linguistics*, Athens, Greece.

Ayers, Gayle. 1996. *Nuclear accent types and prominence: some psycholinguistic experiments*. Ph. D. Dissertation, Linguistics Department, Ohio State University.

Barnes, Jonathan. 2006. *Strength and weakness at the interface*. Belin and New York: Mouton de Gruyter.

Beckman, Jill. 1998. *Positional faithfulness*. Ph. D. Dissertation, University of Massachusetts.

Beckman, Mary E. 1986. *Stress and non-stress accent*. Dordrecht: Foris.

Behrens, S. & S. Blumstein. 1998. Acoustic characteristics of English voiceless fricatives: A descriptive analysis. *Journal of Phonetics* 18: 51-63.

Bertinetto, Pier Marco. 1989. Reflections on the dichotomy 'stress' vs. 'syllable-timing'. *Revue Phonétique appl.*: 99-130.

Bladon, Anthony, Christopher Clark & Katrina Mickey. 1987. Production and perception of sibilant fricatives: Shona data. *Journal of the IPA* 17: 39-65.

Bolinger, Dwight L. 1958. A theory of pitch accent in English. *Word* 14: 109-119.

Bolinger, Dwight L. 1961. Contrastive accent and contrastive stress. *Language* 37: 83-96.

Bolinger, Dwight L. 1972. Accent is predictable (if you're a mind-reader). *Language* 48: 633-644.

Bradlow, Ann R. 1995. A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America* 97 (3): 1916-1924.

Broselow, Ellen, Su-I Chen & Marie K. Huffman. 1997. Syllable weight: convergence of phonology and phonetics. *Phonology* 14: 47-82.

Browman, Catherine. P. & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6: 201-251.

Browman, Catherine P. & Louis Goldstein. 2000. Competing constraints on intergestural coordination and self-organization of phonological strucure. *Les Cahiers de l'ICP, Bulletin de la communication parlé* 5: 25-34.

Byrd, Dani. 1994. *Articulatory timing in English consonant sequences*. Ph. D. Disseration, UCLA.

Byrd, Dani & Eliot Saltzman. 2003. The elastic phase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31: 149-180.

Cambier-Langeveld, Tina & Alice E. Turk. 1999. A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics* 27: 255-280.

Camden, William Greenwood. 1977. *A descriptive dictionary: Bislama to English*. rosebery, Australia: Bridge Printery.

Campbell, Nick & Mary E. Beckman. 1997. Stress, prominence and spectral tilt. In *Intonation: Theory, models and applications* (Proceedings of an ESCA workshop, September 18-20, Athens, Greece), ed. by Antonis Botinis, Georgios Kouroupetroglou & George Carayiannis, 67-70. Athens: ESCA and Dept. of Informatics at University of Athens.

Canfield, D. Lincoln. 1981. *Spanish pronunciation in the Americas*. Chicago: Chicago University Press.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and Topic*, ed. Li, C., 25-55. New York: Academic PRess.

Cho, Taehong 2006. Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. In *Laboratory Phonology 8: Varieties of phonological competence*, eds. Goldstein, L. and D. H. Whalen, 519-548. Berlin and New York: Mouton de Gruyter.

Cho, Taehong, James M. McQueen & Ethan Cox. 2007. Prosodically driven phonetic detals in speech processing: The caseof domain-initial strengthening inEnglish. *Journal of Phonetics* 35: 210-243.

Cho, Taehong, Sun-A. Jun & Peter Ladefoged. 2002. An acoustic and aerodynamic correlates to Korean stops and fricatives. *Journal of Phonetics* 30: 193-228.

Cho, Taehong. 2005. Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realization of /a,i/ in English. *Journal of the Acoustical Society of America* 117: 3867-3878.

Cho, Taehong & P. Keating. 2007. Effects of initial position versus prominence in English. *UCLA working papers in Phonetics* 106: 1-33.

Cho, Taehong & Patricia Keating. 2009. Effects of initial position versus prominence in English. *Journal of Phonetics* 37: 466-485.

Clements, G. N. 1990. The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology I*: 283-333.

Colantoni, Laura. 2001. *Merges, chain shifts and dissimilatory processes: Palatals and rhotics in Argnetinian Spanish*. Ph.D. Dissertation, University of Minnesota.

Colantoni, Laura & Jorge Gurlekian. 2004. Convergence and intonation: historical evidence from Buenos Aires Spanish. *Bilingualism: Language and Cognition* 7: 107-119.

Colantoni, Laura. 2006. Micro and macro sound variation and chcange in Argentinian Spanish. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, eds. Sagarra, N. and A. Jacqueline, 91-102. Somerville, MA: Cascadilla Proceedings Project.

Cole, R. A. & J. Jakimik. 1980. A model of speech perception. In *Perception and production of fluent speech*, ed. Cole, R. A. Hillsdale: Erlbaum.

Chomsky, Noam & Marris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Contreras, Heles. 1977. Spanish Epenthesis and Stress. In *Working papers in Linguistics*, 9-33. University of Washington, Seattle, Washington.

Cooper, Frankilin S., Pierre C. Delattre, Alvin M. Liberman, John M. Borst & Louis J. Gerstman. 1952. Some experiments on the perception of synthetic speech sounds. *Journal of the Acoustical Society of America* 24: 597-606.

Cooper, William E., Stephen J. Eady & Pamela R. Mueller. 1985. Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America* 77: 2142-2156.

Cruttenden, Alan. 1995. Rises in English. In *Studies in general and English phonetics: Essays in honor of professor J. D. O'Connor*, ed. Lewis, J. W., 155-173. London: Routledge.

Cruttenden, Alan. 2006. The deaccenting of given information: A cognitive universal? In *The pragmatic organization of discourse in the languages of Europe*, eds. Bernini, Giuliano and Marcia L. Schwartz, 311-356. The Hague: Mouton de Gruyter.

Crystal, David. 1969. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.

Crystal, Thomas H. & Arthur S. House. 1990. Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustical Society of America* 88: 101-112.

Cutler, Anne & Dennis Norris. 1988. The role of strong syllables in segmentation for lexical access. *journal of Experimental Psychology: Human Perception & Performance* 14: 113-121.

Cutler, Anne, Jacques Mehler, Dennis Norris & Juan Seguí. 1992. The monolingual nature of speech segmentation of French and English. *Cognitive Psychology* 24: 381-410.

Dalcher, Christina. 2006. *Consonant weakening in Florentine Italian: An acoustic study of gradient and variable sound change*. Ph. D disseration, Georgetown University.

Dauer, Rebecca M. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51-62.

Davis, Matthew H., William D. Marslen-Wilsen & M. Gareth Gaskell. 2002. Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *journal of Experimental Psychology: Human Perception & Performance* 28: 218-244.

Davis, Stuart. 1988. Syllable onsets as a factor in stress rules. *Phonology* 5: 1-19.

De Jong, Kenneth, Mary E. Beckman & John Edwards. 1993. The interplay between prosodic structure and coarticulation. *Language and speech* 36: 197-212.

De Jong, Kenneth. 1995. The supraglottal articulation of prominence in English: linguistic stress as localized hperarticulation. *Journal of the Acoustical Society of America* 97: 491-504.

Delattre, Pierre C., Alvin. Liberman & F. Cooper. 1955. Acoustic Loci and Transitional Cues for Consonants. *Journal of the Acoustical Society of America* 27: 769-773.

Delattre, Pierre C. 1964. German phonetics between English and French. *Linguistics* 8: 43-55.

Delattre, Pierre C., A. Liberman & Franklin Cooper. 1964. Formant transitions and loci and as acoustic correlates of place of articulation in American fricatives. *Studia Linguistica* 18: 104-121.

Delattre, Pierre C. 1966. A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics* 4: 183-188.

Delattre, Pierre C. 1969. An acoustic and articulatory study of vowel reduction in four languages. *International Review of Applied Linguistics in Language Teaching* 7: 295-325.

Denes, Paul. 1955. Effect of duration on the perception of voicing. *Journal of the Acoustical Society of America* 27: 761-764.

Dupoux, Emmanuel, Christophe Pallier, Nuria Sebastián & Jacques Mehler. 1997. A destressing "deafness" in French? *Journal of Memory and Language* 36: 406-421.

Eefting, Wieke. 1991. The effect of "information value" and "accentuation" on the duration of Dutch words, syllables, and segments. Journal of the Acoustical Society of America 89 (1): 412-424.

Eddington, David. 2000. Spanish stress assignment within the analogical modeling of language. *Language* 76: 92-109.

Engstrand, Olle & Diana Krull. 2002. Duration of syllable-sized units in casual and elaborated speech: cross-language observations on Swedish and Spanish. *TMH-QPSR* 44: 69-72.

Everett, Daniel & Keren Everett. 1984. Syllable onsets and stress placement in Pirahã. *Proceedings of the West Coast Conference on Formal Linguistics* 3: 105-116.

Fant, Gunnar. 1960. *Acoustic theory of speech production*: Mouton, The Hague.

Fant, Gunnar. 1966. A note on vocal tract size factors and non-uniform F-pattern scalings. *STL-QPSR* 4: 22-30.

Fant, Gunnar. 1975. A non-uniform vowel normalization. In *Technical Report, Speech Transmission Lab. Rep.* Stokholm, Sweden: Royal Inst. Tech.

Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18: 7-44.

Foley, James Addison. 1965. *Spanish Morphology*. Ph. D. Dissertation, MIT.

Forrest, Karen, Gary Weismer, Paul Milenkovic & Ronald N. Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: preliminary data. *Journal of the Acoustical Society of America* 84: 115-123.

Fougeron, Cécile & Patricia Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101: 3728-3740.

Fougeron, Cécile. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29: 109-135.

Fourakis, Marios. 1991. Tempo, stress, and vowel reduction in American English. *The journal of the Acoustical Society of America* 90: 1816-1827.

Fry, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27: 765-768.

Fry, Dennis Butler 1958. Experiments in the perception of stress. *Language and speech* 1: 126-152.

Fudge, Erik C. 1969. Syllables. *JL* 5: 253-287.

Gabriel, Christoph. 2006. Focal pitch accents and subject positions in Spanish: comparing Close-to-Standard varieties and Argentinian Porteño. Paper presented at *Speech Prosody*, Dresden.

Gama Rossi, A. J. 1998. Qual é a natureza do acento secundário no pourtuguês brasileiro? *Cadernos Centro Universitário São Camilo* 4: 77-92.

Gay, Thomas. 1978. Effect of speaking rate on vowel formant movements. *The journal of the Acoustical Society of America* 63: 223-230.

Gili Gaya, Samuel. 1940. La cantidad silâbica en la frase. *Castilla (Valladolid)* 1: 287-298.

Glave, R. D. & A. C. M. Rietveld. 1975. Is the effort dependence of speech loudness explicable on the basis of acoustic cues. *Journal of the Acoustical Society of America* 58: 875-879.

Goedemans, Rob. 1998. *Weightless segments*. The Hague: Holland Academic Graphics.

Gordon, Matthew, Paul Barthmaier & Kathy Sands. 2002. A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetics Association* 32: 1421-1174.

Gordon, Matthew. 2004. A phonological and phonetic study of word-level stress in Chickasaw. *International Journal of American Linguistics* 70: 1-32.

Gordon, Matthew. 2005. A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory* 23: 595-653.

Grice, Martine, Mariapaola D'Imperio, Michelina Savino & Cinzia Avesani. 2005. Strategies for intonation labelling across varieties of Italian. In *Prosodic typology: The phonology of intonation and phrasing*, ed. Jun, S., 362-389. Oxford: Oxford University Press.

Guitart, Jorge. 1978. Aspectos del consonantismo habanero: reexamen descriptivo. *Boletín de la Academia Puertorriqueña de la Lengua Español* 7: 95-114.

Gutiérrez-Díez, Francisco. 2001. The acquisition of English syllable timing by native Spanish speakers learners of English: An empirical study. *IJES* 1: 93-113.

Halle, Morris & Jean-Roger Vergnaud. 1978. *Metrical structures in phonology*. Ms. MIT.

Halliday, Michael A. K. 1967. *Notes on transitivity and theme in English*. London: Longman.

Harris, James W. 1969. *Spanish phonology*. Cambridge, Massachusetts: MIT Press.

Harris, James W. 1975. Stress assignment rules in Spanish. In *1974 Colloquium on Spanish and Portuguese Linguistics*, eds. Milán, W. G., J. J. Staczek and J. C. Zamora. Washington, D. C.: Georgetown University Press.

Harris, James W. 1983. *Syllable structure and stress in Spanish*.vol. 8: Linguistic Inquiry Monograph. Cambridge, Massachusetts: The MIT Press.

Harris, Katherine H. 1958. Cues for discrimination of American English fricatives in spoken syllables. *Language and speech* 1: 1-7.

Harris, M. S. & N. Umeda. 1974. Effect of speaking mode on temporal factors in speech: vowel duration. *Journal of the Acoustical Society of America* 56: 1016-1018.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20: 253-306.

Hedrick, Mark S. & Ralph N. Ohde. 1993. Effect of relative amplitude of frication on perception of place of articulation. *Journal of the Acoustical Society of America* 94: 2005-2026.

Hertz, Susan. 1982. From text to speech with SRS. *Journal of the Acoustical Society of America* 72: 1155-1170.

Hertz, Susan. 1990. A modular approach to multi-language and multi-dialect speech synthesis. *Proceedings of the ESCA Conference on Speech Synthesis.*

Hertz, Susan. 1991. Streams, phones and transitions: toward a new phonological and phonetic model of formant timing. *Journal of Phonetics* 19: 91-109.

Hertz, Susan. 1992. The timing of phones and transitions: a nucleus-based model of English duration. *Working papers of the Cornell Phonetics Laboratory* 5: 135-150.

Hock, Hans Henrich. 1991. *Principles of historical linguistics*. Berlin: Mouton de Gruyter.

Hockett, C. 1958. *A course in modern linguistics*. New York: MacMillan.

Hoequist, Charles. 1983. Durational correlates of linguistic rhythm categories. *Phonetica* 40: 19-31.

Holmquist, Jonathan. 1985. Social correlates of a linguistic variable: A study in a Spanish village. *Language in society* 14: 191-203.

Holmquist, Jonathan. 1998. High Lands - High Vowel: A sample of men's speech in rural Puerto Rico. In *Papers in Sociolinguistics: NWAVE-26*, eds. Paradis, C., D. Vincent, D. Deshaies and M. LaForest, 73-79. Quebec: Université LAval.

Hualde, José Ignacio. 1989. Autosegmental and metrical spreading in the vowel - Harmony systems of northwestern Spain. *Linguistics* 27: 773-805.

Hualde, José Ignacio, Miquel Simonet & Marianna Nadeu. In press. Consonant lenition and phonological recategorization. *Laboratory Phonology.*

Hualde, Jóse Ignacio. 2005. *The sound of Spanish*. New York: Cambridge University Press.

Hubbard, Kathleen. 1995. 'Prenasalised consonants' and syllable timing: evidence from Runyambo and Luganda. *Phonology* 9: 255-265.

Hyman, Larry. 1985. *A theory of phonological weight*. Dordrecht: Foris.

Jassem, Wiktor. 1979. Classification of fricative spectra using statistical discriminant function. In *Frontiers of speech communication*, eds. Lindblom, B. and S. Öhman, 77-91. London: Academic PRess.

Johnson, Keith. 2005. Speaker normalization in speech perception. In *The handbook of speech perception*, 363-389. Oxford: Blackwell Publishers.

Jones, Denes. 1956. *An outline of English phonetics*. Cambridge, England: Cambridge University Press.

Jongman, Allard, Sheila E. Blumstein & Aditi Lahiri. 1985. Acoustic properties for dental and alveolar stops. *Journal of Phonetics* 13: 235-251.

Jongman, Allard, Marios Fourakis & Joan A. Sereno. 1989. The acoustic vowel space of Modern Greek and German. *Language and Speech* 32: 221-248.

Jongman, Allard, Ratree Wayland & Serena Wong. 2000. Acoustic characteristics of American English Fricatives. *Journal of the Acoustical Society of America* 108: 1252-1524.

Jun, Sun-A. 1993. *The phonetics and phonology of Korean prosody*, Ohio State University.

Keating, Patricia & Marie K. Huffman. 1984. Vowel variation in Japanese. *Phonetica* 41: 191-207.

Keating, Patricia, Richard Wright & Jie Zhang. 1999. Word-level asymmetries in consonant articulation. *UCLA working papers in Phonetics* 97: 157-173.

Keating, Patricia, Taehong Cho, Cécile Fougeron & C-S Hsu. 2003. Domain-initial strengthening in four languages. In *Papers in Laboratory Phonology*, 143-161: Cambridge University Press.

Kim, Miran. 2006. Multiple cue interpretation in non-native phoneme perception: Korean listeners' perception of English [s] in varied prosodic contexts. Paper presented at *the 4th Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan*, Honolulu, Hawaii.

Kim, Miran & Hosung Nam. 2009. A comparative study of stress manifestation in Latin American Spanish. Poster presented at *PaPI 09 (Phonetics and Phonology in Iberia)*, Gran Canaria, Spain.

Kim, Miran & Hosung Nam. 2010. Cross-dialectal study on temporal modulation patterns in stress manifestation. Poster presented at *Conference on Laboratory Phonology 12*, Alburquerque, New Mexico.

Klatt, Dennis. 1975. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3: 129-140.

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in English: Acousti and Perceptual evidence. *Journal of the Acoutical Society of America* 59: 1208-1221.

Klatt, Dennis H. 1979. Synthesis by rule of segmental durations in English sentences. In *Frontiers of Speech Communication Research*, eds. Lindblom, B. and S. Öhman. New York: Academic Press.

Kurylowicz, J. 1948. Contribution à la théorie de la syllabe. *BPTJ* 8: 80-114.

Kuzia, Claudia, Taehong Cho & Mirjam Ernestus. 2007. Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics* 35: 301-320.

Ladd, D. Robert. & Nick Campbell. 1991. Theories of prosodic structure: Evidence from syllable duration. *Proceedings of the XII ICPhS*.

Ladefoged, Peter, Jenny Ladefoged & Daniel Everett. 1997. Phonetic structures of Banawá, an endangered langauge. *Phonetica* 54: 94-111.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Cambridge, MA: Blackwell.

Lavoie, Lisa. 2001. *Consonant strength: Phonological Patterns and Phonetic Manifestations*. New York: Garland.

Lavoie, Lisa. 2002. Subphonemic and Suballophonic Consonant Variation: The Role of the Phoneme Inventory. *ZAS Papers in Linguistics* 28: 39-54.

Lehiste, Ilse & G. Peterson. 1961. Transitions, Glides and Diphthongs. *Journal of the Acoustical Society of America* 33: 268-277.

Lehiste, Ilse. 1965. Juncture. In *Proceedings of the 5th International Congress of Phonetic Sciences*, eds. E. Zwirner and W. Bethge, Munster, pp. 172-200.

Lehiste, Ilse. 1970. *Suprasegmentals*: The MIT Press.

Lewis, Anthony. 2001. *Weakening of intervocalic /ptk/ in two Spanish dialects: Toward the quantification of lenition processes.* Ph. D. Dissertation, University of Illinois at Urbana-Champaign.

Liberman, Alvin, Pierre C. Delattre, Franklin Cooper & Louis Gerstman. 1954. The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological monographs* 68: 1-13.

Liberman, Alvin, Pierre C. Delattre, Louis Gerstman & Franklin Cooper. 1956. Tempo of Frequency Change as a Cue for Distinguishing Classes of Speech Sounds. *Journal of Experimental Psychology* 52: 127-137.

Lieberman, Philip. 1960. Some acoustic correlates of word stress in American Engish. *Journal of the Acoustical Society of America* doi: 10.1121/1.1908095. 32: 451-454.

Lindblom, B. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35: 1773-1781.

Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In *Speech production and speech modeling*, eds. Hardcastle, W. J. and A. Marchael. Dordrecht, PP: Kluwer Academic Publishers.

Lindblom, Björn, Augustine Agwuele, Harvey M. Sussman & Elisabet E. Cortes. 2007. The effect of emphatic stress on consonant vowel coarticulation. *Journal of the Acoustical Society of America* 121: 3802-3813.

Llisterri, Joaquim, María Machuca, Carme de la Mota, Contserrat Riera & Antonio Ríos. 2003. The perception of lexical stress in Spanish. Paper presented at *The 15th International Congress of Phonetic Sciences* Barcelona.

Low, Ee, Esther Grabe & Francis Nolan. 2001. Quantitative characterisations of speech rhythm: 'syllable-timing' in Singapore English. *Language and speech* 43: 3777-3401.

Magen, Harriet S. 1997. The extent of vowel-to-vowel coarticulation in English. *Journal of Phonetics* 25: 187-205.

Maiden, Martin. 1991. *Interactive Morphonology: Metaphony in Italy*. New York: Routledge.

Major, Roy C. 1985. Stress and rhythm in Brazilian Portuguese. *Language* 61: 259-282.

Malmberg, Bertil. 1967. The phonetic basis for syllable division. In *Readings in Acoustic Phonetics*, ed. Lehiste, I., 293-300. Cambridge: The MIT Press (Reprinted from Studia Linguistica, Vol. 9, Lund, Sweden: C.W.K. Gleerup, 1955).

Manuel, Sharon Y. & Rena A. Krakow. 1984. Universal and language particular aspects of vowel-to-vowel coarticulation. In *Haskins Lab. Stat. Rep. Speech Res.*, 69-78.

Manuel, Sharon Y. 1990. The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America* 88: 1286-1298.

Marslen-Wilsen, W. D. & P. Zwitserlood. 1989. Accessing spoken words: The importance of word onsets. *journal of Experimental Psychology: Human Perception & Performance* 15: 576-585.

Marslen-Wilson, W. D. & P. Zwitserlood. 1989. Accessing spoken words: the importance of word onsets. *journal of Experimental Psychology: Human Perception & Performance* 15: 576-585.

Mazzaro, Natalia. 2005. Aspiration and velarization of /f/ in Argentine Spanish. *Toronto Working Papers in Linguistics* 25: 58-67.

McCarthy, John. 1984. Theoretical consequences of Montañes vowel harmony. *Linguistic Inquiry* 15: 291-318.

McClean, Michael D. & Stephen M. Tasko. 2002. Association of orofacial with laryngeal and respiratory motor output during speech. *Experimental brain research* 146: 481-489.

McGory, Julia & Manuel Díaz-Campos. 2000. Declarative intonation patterns in multiple varieties of Spanish. Paper presented at *Fourth Hispanic Linguistics Symposium*, Bloomington, Indiana.

Nam, Hosung. 2004. Syllable-level intergestural timing model: split-gesture dynamics focusing on positional asymmetry and moraic structure. In *Papers in Laboratory Phonology*, eds. Cole, J. and J. Hualde.

Nam, Hosung, Louis Goldstein, Eliot Saltzman & Dani Byrd. 2004. TADA: An enhanced, portable Task Dynamics model in MATLAB. *Journal of the Acoustical Society of America* 115: 2430.

Navarro, Tomás. 1948. *El español en Puerto Rico: Contribución a la geografía lingüistica hispanoamericana*: Editorial Universitaria: Universidad de Puerto Rico.

Navarro, Tomás. 1964. La medida de la intensidad. *Boletín del Instituto de Filología de la Universidad de Chile* 16: 231-235.

Nazzi, Thierry, Josiane Bertoncini & Jacques Mehler. 1998. Language discrimination by newborns: toward an understanding o the role of rhythm. *Journal of Experimental Psychology* 24: 756-766.

Nazzi, Thierry & Franck. Ramus. 2003. Perception and acquisition of linguistic rhythm by infants. *Speech Communication* 2003: 233-243.

Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris Publications.

Nittrouer, Susan & Michael Studdert-Kennedy. 1987. The role of coarticulation effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research* 30: 319-329.

Nittrouer, Susan, Michael Studdert-Kennedy & Richard S. McGrowan. 1989. The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research* 32: 120-132.

Nowak, Paweł M. 2006. The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Phonetics* 34: 139-152.

Nooteboom, S. G. 1985. Production and perception of vowel duration: A study of durational properties of vowels in Dutch. Dissertation, Utrecht University, Utrecht, The Netherlands.

O'Connor, Joseph D. & Gordon F. Arnold. 1961. *Intonation of colloquial English*. London: Longmans.

O'Connor, Joseph D. 1973. *Phonetics*. Harmondsworth: Penguin.

Oftedal, Magne. 1985. *Lenition in Celtic and insular Spanish: The second voicing of stops in Gran Canaria*. Oslo: Universitetsforlaget cop.

Öhman, S. E. G. 1966. Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America* 39: 151-168.

Olive, Joseph P., Alice Greenwood & John S. Coleman. 1993. *Acoustics of American English Speech*. New York: Springer Verlag.

Olsen, L. 1972. Rhythmical patterns and syllablic features of the Spanish sense group. Paper presented at *The 7th International Congress of Phonetics Sciences*, Montreal (1971).

Ortega-Llebaria, Marta & Pilar Prieto. 2005. Stress and focus in Spanish and Catalan: Patterns of duration and vowel quality. Paper presented at *Phonetics and Phonology in Iberia*, Barcelona, Spain.

Ortega-Llebaria, Maria. 2006. Cues to Stress and Accent in Spanish. *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology*: 104-118.

Ortega-Llebaria, Maria & Pilar Prieto. 2006. Stress and accent in Catalan and Spanish. Paper presented at *Speech Prosody*, Dresden, Germany.

Ortega-Llebaria, Marta. 2008. Comparing the 'Magnifying Lens' effect of stress to that of contrastive focus in Spanish. Paper presented at *the 3rd Conference on Laboratory Approaches to Spanish Phonology*, Somerville, MA.

Ortega-Llebaria, Marta & Pilar Prieto. 2011. Acoustic correlates of stress in Central Catalan and Castilian Spanish. *Language and Speech* 54 (1): 73-97.

Palmer, Harold E. 1922. *English intonation with systematic exercises*. Cambridge: Heffer.

Parmenter, C. E. & S. N. Treviño. 1935. The length of the sounds of a Middle Westerner. *American Speech* 10: 129-133.

Partridge, Monica A. 1950. The influence of stress on vowel quality in spoken Russian. *The Slavonic and East European Review* 29: 244-255.

Peterson, Gordon E. 1955. An oral communication model. *Language* 31: 414-427.

Peterson, Gordon E. & Harold L. Barney. 2002. Non-uniform scaling based on speaker normalization. Paper presented at *In Proceedings of the International Conference on Acoustic, Speech and Signal Processing (ICASSP)*.

Peterson, Gordon E. & Ilse Lehiste. 1960. Duration of Syllable nuclei in English. *Journal of the Acoustical Society of America* 32: 693-703.

Pierrehumbert, Janet & David Talkin. 1992. Lenition of /h/ and glottal stop. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, eds. Docherty, G. and D. R. Ladd, 90-117. Cambridge: Cambridge University Press.

Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English intonation*. Ph. D. disser0tation, MIT.

Pike, Kenneth L. & Eunice V. Pike. 1947. Immediate constituents of Mazatec syllables. *IJAL* 13: 78-91.

Pike, Kenneth L. 1945. The intonation of American English. In *Intonation*, ed. Bolinger, D. L., 53-83. Harmondsworth: Penguin.

Port, Robert F. & Jonathan Dalby. 1982. Consonant/Vowel ratio as a cue for voicing in English. *Perception & Psychophysics* 32: 142-152.

Prince, Alan S. & Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Ms. Rutgers University and University of Colorado at Boulderm.

Quené, H. 1989. *The influence of acoustic-phonetic word boundary markers on perceived word segmentation in Dutch*. Dissertation, Utrecht University, Utrecht, The Netherlands.

Quilis, Antonio & Manuel Esgueva. 1983. 'Realizacion de los fonemas vocalicos espanoles en posicion fonetica normal'. *Estudios de fonetica* 1: 159-252.

Quilis, Antonio. 1993. *Tratado de Fonética y Fonología Españolas*. Madrid: Editorial Gredos.

Roach, Peter. 1982. On the distinction between 'stress-timed' and 'syllable-timed' langauges. In *Linguistic controversies: Essays in linguistic theory and practice in honor of F. R. Palmer*, ed. Crystal, T. H., 73-79. London: Edward Arnold.

Salverda, Anne P., Delphine Dahan & James M. McQueen. 2003. The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension: Evaluating a linking

hypothesis between fixation and linguistic processing. *Journal of Psycholinguistic Research* 29: 557-580.

Schmid, S. 2004. Une approche phonétique de l'isochronie dans quelques dialectes italo-romans. In *Nouveaux départs en phonologie*, eds. Meisenburg, T., Selig, M., Tübingen: G. Narr, 109-124.

Selkirk, Elisabeth O. 1978. On prosodi structure and its relation to syntactic structure. In *Nordic Prosody II*, ed. Fretheim, T., 111-140. Trondheim: Tapir.

Selkirk, Elisabeth O. 1982. Syllables. In *The structure of phonological representations 2*, eds. Hulst, van der and N. Smith, 337-383. Dordrecht: Foris.

Shadle, Christine H. & Sheila J. Mair. 1996. Quantifying spectral characteristics of fricatives. Paper presented at *ICSLP*.

Shatz, Carole D. 1954. The role of context in the perception of stops. *Language* 30: 47-56.

Simpson, Adrian P. 2001. Dynamic consequences of differences in male and femal vocal tract dimensions. In *Proceeding of 6th International seminar on speech production*. Sydney.

Sluijter, Agaath. M. C. 1995. *Phonetic correlates of stress and accent*. The Hague: Holland Academic Graphics.

Sluijter, Agaath. M. C. & Vincent J. van Heuven. 1996a. Acoustic correlates of linguistic stress and accent in Dutch and American English. Paper presented at *Fourth international conference on spoken language processing*, Philadelphia.

Sluijter, Agaath. M. C. & Vincent J. van Heuven. 1996b. Spectral balance as an acoustic correlate of linguistic sress. *Journal of the Acoustical Society of America* 100 (4): 2471-2485.

Sluijter, Agaath M. C., Vincent J. van Heuven & Jos J. A. Pacilly. 1997. Spectral balance as a cue in the perception of linguistic sress. *Journal of the Acoustical Society of America* 101 (1): 503-513.

Smith, Jennifer. 2002. *Phonological augmentation in prominent positions*. Ph.D. dissertation, University of Massachusetts, Amherst.

Solan, L. 1979. Spanish Adverbs: Support for the Phonological Cycle? In *University of Massachusetts Working Papers in Linguistics*, ed. Lowenstamm, J., 602-605. Amherst, Massachusetts.

Solé, Maria-Josep S. 1991. Stress and rhythm in English. *Revista Alicantina de Estudios Ingleses* 4: 145-162.

Stevens, Kenneth N. 1999. *Acoustic phonetics*. Cambridge: MIT Press.

Stevens, Kenneth N. 1972. The quantal nature of speech: evidence from articulatory-acoustic data. In *Human communication: A unified view*, eds. Denes, P.B. and E.E. David Jr., 51-61. New York: McGraw Hill.

Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17: 3-46.

Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge, Massachusetts: The MIT Press.

Stevens, Mary E. *A phonetic investgation into Raddoppiamento Sintattico in Sienese Italian speech*. Ph. D Dissertation, School of Languages and Linguistics, The University of Melbourne.

Stevens, Peter. 1960. Spectra of fricative noise in human speech. *Language and speech* 3.

Suomi, Kari & Riikka Ylitalo. 2004. On durational correlates of word stress in Finnish. *Journal of Phonetics* 32: 35-63.

Tabain, Marija, Guillaume Rolland & Christophe Savariaux. 2001. Coarticulatory effects at prosodic boundaries: some acoustic results. Paper presented at *Euro Speech 2001*.

Tiffany, William R. 1959. Nonrandom sources of variation in vowel quality. *Journal of Speech and Hearing Research* 2: 305-317.

Trubetskoy, N. S. 1979 [1938]. *Principles de Phonologie*. Klinksieck: Paris.

Turk, Alice E. & James R. Sawusch. 1996. The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America* 99: 3782-3790.

Turk, Alice E. & James R. Sawush. 1997. The domain of accentual lengthening in American English. *Journal of Phonetics* 25: 24-41.

Turk, Alice E. & Laurence White. 1999. Structural influences on accentual lengthening. *Journal of Phonetics* 27: 171-206.

Turk, Alice E. & Stefanie Shattuck-Hufnagel. 2000. Word-boundary-related durational patterns in English. *Journal of Phonetics* 25: 25-41.

Tyler, Michael D. & Anne Cutler. 2009. Cross-language differences in cue use for speech segmentation. *Journal of the Acoustical Society of America* 126: 367-376.

Umeda, Noriko. 1977. Consonant duration in American English. *Journal of the Acoustical Society of America* 61: 846-858.

Vaissière, Jacqueline. 1988. Prediction of velum movement from phonological specifications. *Phonetica* 45: 122-139.

van der Hulst, Harry. 1999. *Word prosodic systems in the languages of Europe*. Berlin: Mouton de Gruyter.

Vennemann, Theo. 1988. *Preference laws for syllable structure and the explanation of sound change: With special reference to German, Germanic, Italian, and Latin.* Berlin: Mouton de Gruyter.

Verner, Karl. 1877. Eine ausnahme der ersten lautverschiebung. *Kuhn's Zeitschrift (KZ)* 23: 97-130.

Wagner, Michael. 2006. Givenness and Locality. In *Proceedings of SALT 16*, eds. Howell, Jonathan and Masayuki Gibson. Ithaca: CLC Publications.

Whalen, D. H. 1981. Effect of vocalic formant transitions and vowel quality on the English [s] - [sh] boundary. *Journal of the Acoustical Society of America* 69: 275-282.

Whalen, Douglas H. 1991. Perception of the English /s/-/sh/ distinction relies on fricative noises and transitions, not on brief spectral slices. *Journal of the Acoustical Society of America* 90: 1776-1785.

White, Laurence & Sven L. Matthys. 2007. Calibrating rhythm: first language and second language studies. *Journal of Phonetics* 35: 501-522.

White, Laurence, Elinor Payne & Sven L. Mattys. 2009. Rhythmic and prosodic contrast in Venetan and Sicilian Italian. In *Phonetics and Pho nology: Interactions and Interrelations*, eds. M. Vigario, S. Frota & M. J. Freitas, Amsterdam: John Benjamins. 137-155.

Whitley, Stanley. 1976. Stress in Spanish: Two approaches. *Lingua* 39: 301-332.

Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91: 1707-1717.

Zwicker, Eberhard & Richard Feldtkeller. 1967. *Das Ohr als Nachrichtenempfanger*. Stuttgart: Hirzel.

Appendix

A- 1 Complete word list

Spanish Data			Stops						Affricate	Fricatives			
PoW	stress	V	/p/	/b/	/t/	/d/	/k/	/g/	/tʃ/	/f/	/s/	/ʝ/	/x/
ini	str	i	pícaro	bígaro	títere	díselo	quítalo	guísalo	chícharo	fíchalo	sísalo	Yínato	jícara
	unstr	i	picota	bisagra	tirado	dibujo	quilate	guisado	chicharra	fichero	silicio	Yimoto	jinete
	str	u	púdico	búfalo	túnica	dúdalo	cúmulo	Gúdico	chútalo	fúgate	súbito	lluvia	júbilo
	unstr	u	pupila	bufete	tugurio	dudoso	culebra	gusano	chupado	futuro	sujeto	lluvioso	jugoso
	str	e	pésalo	bélico	tétrico	déjelo	quédate	guésalo	chévere	féretro	séquito	llévelo	jefazo
	unstr	e	pesado	bellaco	tesina	derroche	quejica	guerrilla	chequera	fetiche	secado	llegada	jefazo
	str	o	póliza	bóveda	tópico	dóciles	kósovo	gótico	chócalo	fósiles	sótano	llóralo	jódelo
	unstr	o	pollera	botones	toquilla	dotado	cosecha	goloso	chorizo	forofo	sollozo	lloroso	jojoba
	str	a	pábulo	báculo	támesis	dádiva	cábala	gánalo	cháchara	fábrica	sádico	llámalo	jálalo
	unstr	a	pachucho	bajada	tapujo	dativo	cabeza	garrafa	chabola	fatiga	saludo	llavero	jarabe
med	str	i	pupila	ubique	fatiga	redicho	toquilla	maguito	rechifla	sufijo	risible	tullido	quejica
	unstr	i	tópico	súbito	gótico	dádiva	séquito	lánguido	achicar	gráfico	fósiles	bullicioso	lógico
	str	u	tapujo	dibujo	futuro	deduje	locura	tugurio	pachucho	refugio	desuso	velludo	lujuria
	unstr	u	púrpura	pábulo	rótulo	cédula	báculo	lúgubre	achuchar	prófugo	cápsula	velludito	brújula
	str	e	ropero	cabeza	litera	ladera	chequera	reguero	fichero	bufete	cosecha	pollera	sujeto
	unstr	e	áspero	bébelo	títere	jódelo	búsqueda	páguelo	fíchelo	rífelo	díselo	óyelo	déjelo
	str	o	reposo	chabola	botones	dudoso	picota	jugoso	lechoso	aforo	pasota	sollozo	jojoba
	unstr	o	reposar	símbolo	tórtola	ídolo	lógico	prólogo	pachucho	forofo	kósovo	barullo	badajo
	str	a	chupado	jabato	dotado	badajo	secado	llegada	chicharra	jefazo	bisagra	bellaco	bajada
	unstr	a	párpado	cábala	sótano	quédate	chócalo	bígaro	cháchara	ráfaga	pésalo	óyalo	pájaro

Note: Spanish orthography uses ' ' ' for initial and final stresses but leaves penultimate unmarked.

A- 2 Within-dialect means of F1, F2, and F3 by stress and gender

(a) Male		/i/			/u/			/e/			/o/			/a/		
		F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
AS	Str	307	2120	2572	339	1093	2620	415	1930	2551	477	1193	2652	665	1473	2458
	SD	24	302	261	24	102	155	26	94	200	60	84	150	67	152	129
	Unstr	329	2156	2587	377	1309	2546	390	1791	2519	436	1273	2587	577	1509	2496
	SD	42	204	220	38	215	182	41	109	162	40	214	154	65	114	119
PS	Str	310	2119	2742	346	1038	2615	415	1855	2493	445	998	2635	617	1413	2494
	SD	48	76	322	42	135	51	62	120	123	70	113	198	63	93	227
	Unstr	337	1994	2512	347	1189	2590	417	1813	2500	425	1205	2582	537	1401	2487
	SD	48	122	206	61	241	240	72	159	335	73	317	289	89	108	119

(b) Female		/i/			/u/			/e/			/o/			/a/		
		F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
Spanish	Str	364	2671	3221	413	1177	2929	487	2234	2881	546	1251	2878	876	1704	2772
	SD	27	62	143	33	97	114	35	67	106	37	42	74	55	158	171
	Unstr	379	2490	3121	432	1424	2910	480	2187	2924	520	1336	2879	773	1800	2871
	SD	50	203	287	42	289	85	45	82	86	69	102	179	65	165	140
Peruvian Sp.	Str	346	2718	3301	407	1258	3628	557	2324	3266	593	1336	3502	920	1841	3159
	SD	47	171	111	32	148	169	118	222	108	51	49	174	36	112	189
	Unstr	405	2484	3351	413	1425	3572	563	2167	3345	521	1500	3470	740	1835	3362
	SD	42	148	109	52	274	184	117	188	181	99	280	190	76	161	182

A- 3 Summary of correlation table

Pearson Correlation result

		Dialect	GEN	PoW	STR	Cincl	S_RMS	Vincl	V_RMS	SYLL
Dialect	R	1	-.106*	0	0	.111**	.141**	-0.014	.214**	0.065
	p		0.014	1	1	0.01	0.001	0.742	0	0.132
	N	600	600	600	600	600	600	600	600	600
GEN	R	-.106*	1	0	0	0.069	-0.023	-0.058	.266**	0.01
	p	0.014		1	1	0.108	0.594	0.179	0	0.821
	N	600	600	600	600	600	600	600	600	600
PoW	R	0	0	1	0	-.255**	0.017	0.013	-0.02	-.162**
	p	1	1		1	0	0.686	0.763	0.642	0
	N	600	600	600	600	600	600	600	600	600
STR	R	0	0	0	1	-.464**	0.032	-.609**	-.198**	-.689**
	p	1	1	1		0	0.458	0	0	0
	N	600	600	600	600	600	600	600	600	600
Cincl	R	.111**	0.069	-.255**	-.464**	1	-0.004	.204**	.120**	.792**
	p	0.01	0.108	0	0		0.929	0	0.005	0
	N	600	600	600	600	600	600	600	600	600
S_RMS	R	.141**	-0.023	0.017	0.032	-0.004	1	.166**	.458**	.101*
	p	0.001	0.594	0.686	0.458	0.929		0	0	0.019
	N	600	600	600	600	600	600	600	600	600
Vincl	R	-0.014	-0.058	0.013	-.609**	.204**	.166**	1	.246**	.759**
	p	0.742	0.179	0.763	0	0	0		0	0
	N	600	600	600	600	600	600	600	600	600
V_RMS	R	.214**	.266**	-0.02	-.198**	.120**	.458**	.246**	1	.233**
	p	0	0	0.642	0	0.005	0	0		0
	N	600	600	600	600	600	600	600	600	600
SYLL	R	0.065	0.01	-.162**	-.689**	.792**	.101*	.759**	.233**	1
	p	0.132	0.821	0	0	0	0.019	0	0	
	N	600	600	600	600	600	600	600	600	600

** Correlation is significant at the 0.01 level (2-tailed)

*. Correlation is significant at the 0.05 level (2-tailed)

A- 4 Gender effect summary: Repeated Measures ANOVA

(a) Tests of Within-Subjects Contrasts

Source	Measure	PoW	Stress	Type III Sum of Squares	df	Mean Square	F	Sig.
PoW * GEN	Cdur	Linear	Stress	101.040	1	101.040	.799	.395
	Vdur	Linear	Stress	21.769	1	21.769	.603	.457
	Sylldur	Linear	Stress	29.010	1	29.010	.140	.717
Error(PoW)	Cdur	Linear	Stress	1137.530	9	126.392		
	Vdur	Linear	Stress	325.119	9	36.124		
	Sylldur	Linear	Stress	1868.975	9	207.664		
Stress * GEN	Cdur	PoW * Stress	Linear	78.539	1	78.539	1.108	.320
	Vdur	PoW * Stress	Linear	11.435	1	11.435	.282	.609
	Sylldur	PoW * Stress	Linear	149.909	1	149.909	1.613	.236
Error(Stress)	Cdur	PoW * Stress	Linear	637.879	9	70.875		
	Vdur	PoW * Stress	Linear	365.510	9	40.612		
	Sylldur	PoW * Stress	Linear	836.392	9	92.932		
PoW * Stress * GEN	Cdur	Linear	Linear	118.080	1	118.080	1.358	.274
	Vdur	Linear	Linear	48.001	1	48.001	2.817	.128
	Sylldur	Linear	Linear	15.509	1	15.509	.134	.722
Error (PoW*Stress)	Cdur	Linear	Linear	782.463	9	86.940		
	Vdur	Linear	Linear	153.374	9	17.042		
	Sylldur	Linear	Linear	1038.133	9	115.348		

(b) Tests of Between-Subjects effects

Source	Measure	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Cdur	535721.714	1	535721.714	1359.765	.000
	Vdur	230119.572	1	230119.572	1117.046	.000
	Sylldur	1.468E+06	1	1.468E+06	2597.218	.000
GEN	Cdur	968.777	1	968.777	2.459	.151
	Vdur	82.244	1	82.244	.399	.543
	Sylldur	1615.561	1	1615.561	2.858	.125
Error	Cdur	3545.831	9	393.981		
	Vdur	1854.065	9	206.007		
	Sylldur	5087.213	9	565.246		

A- 5 RMS ratio summary: C-ss/V-ss ratio in /s/V syllables

		RMS (unit in Pa)				RMS ratio (C/V)	
Dialect		Argentinian Sp.		Peruvian Sp.		Argentinian Sp.	Peruvian Sp.
Vowel	Stress	C_RMS	V_RMS	C_RMS	V_RMS	Mean	Mean
/i/	Str	0.0145	0.0387	0.0159	0.0482	**0.39**	**0.34**
	Unstr	0.0154	0.0395	0.0179	0.0452	**0.43**	**0.42**
/u/	Str	0.0137	0.0414	0.0142	0.0501	**0.34**	**0.29**
	Unstr	0.0109	0.0343	0.0154	0.0375	**0.35**	**0.48**
/e/	Str	0.0161	0.0454	0.0159	0.0565	**0.36**	**0.29**
	Unstr	0.0146	0.0393	0.0200	0.0475	**0.40**	**0.42**
/o/	Str	0.0152	0.0423	0.0161	0.0547	**0.38**	**0.30**
	Unstr	0.0149	0.0373	0.0170	0.0460	**0.40**	**0.38**
/a/	Str	0.0153	0.0531	0.0175	0.0553	**0.33**	**0.34**
	Unstr	0.0153	0.0415	0.0175	0.0479	**0.42**	**0.37**

*Note: (a) the total N=280, each sub group N=28. (b) C_RMS and V_RMS refer to RMS values of onset and vowel in /s/V syllables. In this RMS measurement, transitions are not included.

A- 6 Spectral balance summary for all vowels: Repeated Measures ANOVA [unaccented condition]

Source='stress'		Tests of Within-Subjects Contrasts					
Vowel	Measure	Type III Sum of Squares	df	Error	Mean Square	F	Sig.
/i/	DB	2.09	1	9	2.09	2.097	0.182
	B1	2.313	1	9	2.313	2.398	0.156
	B2	15.124	1	9	15.124	2.82	0.127
	B3	54.406	1	9	54.406	25.937	0.001
	B4	30.981	1	9	30.981	5.952	0.037
/u/	db	28.595	1	9	28.595	27.268	0.001
	b1	32.637	1	9	32.637	27.664	0.001
	b2	7.05	1	9	7.05	2.604	0.141
	b3	0.181	1	9	0.181	0.077	0.787
	b4	0.596	1	9	0.596	0.177	0.684
/e/	db	18.352	1	9	18.352	15.244	0.004
	b1	8.483	1	9	8.483	5.76	0.04
	b2	42.859	1	9	42.859	13.642	0.005
	b3	75.257	1	9	75.257	16.112	0.003
	b4	200.16	1	9	200.16	33.736	0
/o/	db	27.285	1	9	27.285	14.166	0.004
	b1	5.604	1	9	5.604	3.734	0.085
	b2	98.699	1	9	98.699	26.794	0.001
	b3	187.887	1	9	187.887	72.731	0
	b4	166.443	1	9	166.443	32.542	0
/a/	db	21.519	1	9	21.519	9.627	0.013
	b1	0.235	1	9	0.235	0.165	0.694
	b2	35.901	1	9	35.901	8.642	0.016
	b3	137.364	1	9	137.364	48.951	0
	b4	47.513	1	9	47.513	8.297	0.018