The Pennsylvania State University

The Graduate School

Department of Spanish, Italian and Portuguese

AN EXPERIMENTAL APPROACH TO SYLLABLE WEIGHT AND STRESS IN SPANISH

A Thesis in

Spanish

by

Michael Shelton

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2007

The thesis of Michael Shelton was reviewed and approved* by the following:

Henry J. Gerfen
Associate Professor of Spanish and Linguistics
Thesis Advisor
Chair of Committee

Judith F. Kroll
Liberal Arts Research Professor of Psychology and Linguistics

John M. Lipski
Edwin Erle Sparks Professor of Spanish and Linguistics

B. Richard Page
Associate Professor of German and Linguistics

Nicolás Gutiérrez Palma
Assistant Professor of Psychology – University of Jaén, Spain
Special Member

William R. Blue
Professor of Spanish
Head of the Department of Spanish, Italian and Portuguese

*Signatures are on file in the Graduate School

ABSTRACT

This work examines the cognitive representation of phonotactic constraints on Spanish stress via the collection of behavioral data. Critical stimuli in four experiments consist of nonwords that violate the Spanish stress window. Experiment 1 finds statistical differences among latency and accuracy data for stimuli that represent theoretically proscribed sequences, theoretically licit sequences that are unattested for diachronic reasons, and fully licit gaps. Experiments 2 and 3 find differential patterning of rising and falling diphthongs. Experiment 4 tests the time course of phonological encoding and evidences differential treatment of both diphthong types across delays in a delayed naming task. The data are interpreted as evidence in favor of sensitivity to a continuum of weight across the lexicon. The findings challenge traditional approaches to syllable weight which call for a categorical, binary light/heavy distinction. Rather, the results accord with a stochastic or probabilistic conception of the lexicon in which speakers of a language are able to track statistical patterns of extant combinations of sounds in order construct a phonological grammar that is gradient in nature. Implications for two prominent psycholinguistic models of speech production (Levelt et al., 1999; Dell 1986, 1988) are also discussed. Specifically, the findings challenge the proponents of each model to account better for the interaction of stress and syllable-level encoding, both as represented in the lexicon and across time during processing.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

It is impossible to sum up in a couple pages how grateful I am to all the people who have supported me throughout this endeavor. Nonetheless, I would like to take a moment to acknowledge at least a few of the people who have helped me along the way, and to whom I offer my most heartfelt thanks.

First and absolutely foremost, I owe a huge debt of gratitude to my advisor, Chip. You kept me on track throughout my years here at Penn State. Whether it was in our little closet of a lab space in Mexico, here on campus, or during my "dark night of the soul," you were always there to give me the push I needed. Thank you.

To my committee, Judy, John, Richard, and Nico, I am also indebted for all of your guidance. Judy, thank you for being my inspiration to explore psycholinguistics. It was in your class that I first got a taste of what would end up being half of my dissertation! John, while my research did not follow the sociolinguistic path, I enjoyed every course I took with you, and your insights on Hispanic diachronic linguistics were a needed support for the end result of this dissertation. Richard, danke für die Hilfe, und dass Du mir immer den Rücken gestärkt hast. Your comments as an outside member pushed me to think about interesting questions from a perspective I had not previously considered. And, Nico, muchísimas gracias por tanto apoyo. Sin ti nunca hubiera podido pasar tantos participantes, ni entendería los modelos de producción del habla. Me invitaste a tu casa durante mi tiempo en Jaén, y sin tu ayuda, ¡me hubiera perdido completamente entre tantos análisis estadísticos! Te lo agradezco de corazón.

For help in Puebla, I would like to thank Olivia Quiroz at the Iberoamerican University, my *pucelana* friend Cristina Bayón, and Tere Sosa de Monterrosas (mi madre mexicana). For help here at Penn State, I am indebted to Wendy Rizzo and Gabriela Salerno (more than you know, thanks so much to you both).

Also at Penn State I would like to thank the Center for Language Science. I could not envision a better group of language scientists. The intrinsic nature of the group offers a forum for discussion of work from such different perspectives. You all helped me to focus my work better than any single advisor could. I'd also like to thank Becky, Lindsey, Carol, and Laurie in the office. Laurie, our chats kept me sane.

Two other professors I will never forget are Giuli and Marie. Giuli, your psycholinguistic insights, as well as smiles and hugs, were invaluable. Gracias de verdad. And, Marie, your history of the language course is quite possibly the best course I have taken, ever. You have been a true inspiration to so many students who have walked the halls of Burrowes. I am no exception.

A big hug to all of my friends and colleagues in "the cubes."

I would also like to thank the College of Liberal Arts and the Department of Spanish, Italian and Portuguese at Penn State for the Sparks Fellowship, the Humanities Initiative Dissertation Support, two Research and Graduate Studies Office (RGSO) support grants, and travel grants. Without such support, it would have been impossible for me to finish my work on time, and I definitely would be much farther in debt!

Outside of Penn State, I would like to thank Jim O'Neill, Elaine Fuller Carter, Phyllis VanBuren, Isolde Müller, and Shawn Jarvis for pushing me towards graduate school in the first place.

To my family, you know I love you guys. Thank you for supporting me from day one.

Lastly, to my grandfather. I wish you could be here to see this.

Chapter 1

Introduction and Background

1.1 Introduction

One of the long-standing conundrums in the linguistic study of the Spanish language is how to account for its stress system. The task is particularly difficult, given the conflicting issues of inherited influence from Latin and the evolution of Spanish away from Latin stress patterns. There is no consensus in the literature on how the stress system actually functions. A central part of the dispute revolves around whether Spanish stress is sensitive to the concept of syllable weight, and all possible conclusions are controversial. Traditional analyses in the literature fail to provide a means of deciding between options, leaving the puzzle unresolved. In this dissertation, I propose to extend previous approaches through the collection and analysis of behavioral data. Such data, I argue, provide a window into the cognitive representation of the phonology of Spanish. Traditional approaches have been able to describe extant patterns in the Spanish lexicon. In so doing, they have hypothesized phonotactic constraints that restrict the presence of certain sound sequences in the language. Historical linguists have identified patterns in the language that are predicted to be absent due to restraints on syllable structure in Latin. However, without behavioral data, it is difficult, if not impossible, to ascertain how these patterns are represented cognitively by native speakers.

While traditional approaches to linguistic scholarship have amassed a large body of research that has provided great insight into the structures of human languages, a fundamental drawback is that such approaches are often disconnected from theories of the mind and the cognitive representation and processing of language. The present work is situated within the growing area of experimental linguistics. Specifically, I adapt behavioral techniques employed in cognitive psychology to examine topics of interest in Spanish linguistics. This dissertation seeks to overcome the limitations of traditional linguistic approaches by providing novel analyses of what has been a long-standing, recalcitrant problem in Hispanic linguistics. In a broader sense, this research will bring new data to the field in order to refine our understanding of language and the brain, to test between competing linguistic hypotheses and theories, and to develop more integrated models of language structure and language processing.

The remainder of this chapter is organized as follows. In §2 I discuss traditional and emergentist approaches to phonotactics and phonology. In §3 I describe the problematic nature of Spanish stress and its relation to syllable structure. §4 reviews models of reading and of speech production in the psycholinguistic literature and identifies the methodology employed to test lexical processing. Finally, in §5 I outline the structure of the remaining chapters of this dissertation.

1.2 Phonology and phonotactic knowledge

Native speakers of a language have remarkably consistent intuitions about how sounds may be combined to form words in their language. English, for example, permits the consonant cluster /nd/ in words such as *and*. However, predictably all native speakers of English will agree that the repositioning of the consonants at the beginning of a sequence of segments, such as in *nda*, produces a combination of sounds that is impossible in their language. Speakers of other languages do not share this intuition, as many Bantu and Mesoamerican languages, for example, allow prenasalization of initial stops. In addition to clusters, there are also restrictions in English on single consonants. Any native speaker of English will accept the velar nasal /ŋ/ at the end of syllables, such as in the word *king* /kIŋ/ but will readily reject the same segment in initial position: */ŋIk/. Such a restriction on segmental positioning within words is not a linguistic universal; numerous languages, such as Welsh or Vietnamese, evidence initial velar nasal consonants. The observation of language particular patterns such as these has long been used to motivate phonology, the focus of which has long been the characterization of the systematic nature of sound organization both within and across languages. These specific phonotactic constraints, restrictions on the environments in which sounds may or may not appear, are part of what defines the phonological system of English.

1.2.1 Generative approaches to phonotactic encoding

While decades of research have described the phonological patterns of the languages of the world (e.g. Ladefoged & Maddieson, 1996, on sound inventories in the world's languages), it is not clear exactly how phonotactic constraints are encoded by the language processing system. Different schools of thought draw from distinct assumptions as to the architecture of the mental lexicon. The main tenet of transformational-generative linguistics is the existence of a universal grammar, i.e. "the a priori *faculté de langage*" shared by all human beings that allows language acquisition to be possible (Chomsky & Halle, 1968, p. 4). It is these linguistic universals that determine the structure of the specific grammars of languages, and the form and organization of phonological rules. Linguists of the traditional generative school view a linguistic grammar as a set of rules which are applied to an abstract underlying representation. The application of a phonological rule creates the phonetic forms that underlie their realization in speech. The motivation for this belief is the hypothesis that the human capacity for language is designed in such a way as to minimize the amount of information that must be stored in the speaker's mental lexicon. All predictable information is stored as grammatical rules and principles, and distinctive information is stored symbolically in the lexicon.[1]

There is still debate, however, as to what levels of representation are used to encode phonotactic knowledge (Goldrick, 2004). Early generative work, such

[1] For further discussion on the specificity of phonological representation, see Steriade's work on lexical conservatism, e.g. Steriade, 1997a, 1997b, 1999a.

as the seminal *The Sound Pattern of English* (Chomsky & Halle, 1968), considers the segment the basic level of phonological encoding. However, more recent arguments propose that phonotactic information may be stored at the featural or syllabic levels. Research on the phonetics-phonology interface, such as Archangeli & Pulleyblank's (1994) theory of grounded phonology, argues for a phonetically motivated account of phonotactic constraints. In phonetics-based theories, the patterns observed in spoken language are a reflection of the natural tendency towards articulatory ease and increased perceptual contrast. As the generative perspective asserts a lexicon in which only information that is not predictable is encoded, then one may assume that the encoding of phonotactic elements at the segmental level would be redundant, given the predictability of certain surface-level processes due to phonetic universals. Phonetics-based phonological theories, therefore, argue for underspecification at the symbolic level of the phonological segment (Archangeli & Pulleyblank, 1994; see also Hayes et al., 2004).

Other researchers have highlighted the importance of the syllable in the processing of phonotactic rules. The syllable can be a problematic construct in phonological theory, as it lacks any direct phonetic correlates and requires even further levels of abstraction. Nevertheless, various researchers have adopted phonotactic patterns as motivation for syllable structure. Kahn (1976) reintroduced the concept of the syllable to generative phonology after Chomsky & Halle (1968). Kahn's theory of the syllable consists of a hierarchical model that connects the segments of the underlying representation to syllable nodes above.

In the example below ("Jennifer"), we can see the theory is eloquent in that it renders a clear visual representation of ambisyllabicity, mirroring English speakers' inconsistent intuitions on syllable boundaries:[2, 3]

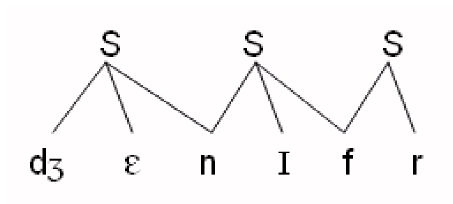

Figure **1**: Ambisyllabicity in English according to Kahn (1976).

Clements & Keyser (1983) expanded on Kahn's (1976) thesis by exploring the internal structure of the syllable. They posit a three-tiered theory of syllable structure in which the syllable node and segments are connected through a CV tier which marks syllabicity by identifying each segment as a syllable peak (V) or syllable margin (C)[4].

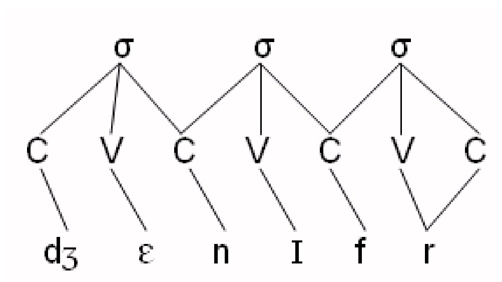

Figure **2**: Ambisyllabicity in English according to Clements & Keyser (1983).

---

[2] This same inconsistency is taken by Steriade (1999b) as evidence disfavoring a syllable-based account of phonotactic knowledge, arguing that syllable boundaries cannot be read in the auditory stimulus.
[3] Figure adapted from Clements & Keyser (1983).
[4] Figure adapted from Clements & Keyser (1983).

Identifying phonotactic gaps such as the trisyllabic consonant clusters *tpt, dbd, tkt, dgd*, Kahn's (1976) model is also able to account for nonexistent patterns in English by delineating constraints on the distribution of consonants in English syllables. He proposes three hypotheses whose combination explains this phonotactic gap in English: English onsets cannot contain stop-stop sequences (hence **Vt.ptV*), alveolar stops cannot precede non-coronals in English codas (hence **Vtp.tV*), and each segment/feature must belong to some syllable (hence **Vt.p.tV*) (Kahn, 1976, as described in Steriade, 1999). Crucially, it is not just English as a language that bans these sequences. We need the concept of the syllable to understand under precisely what circumstances such sequences can arise.

In addition to representation and gaps, phonological processes also render support for a syllable level of encoding. Kenstowicz (1994) cites epenthesis/deletion strategies in words like *rhythm* (epenthetic schwa, [rɪðə̠m]) and *damn* (deleted /n/, [dæm]) as illustrative of the importance of syllable structure. It is difficult to explain why languages would insert vowels or delete consonants out of nowhere, unless we take into consideration phonotactic constraints at the syllable level, i.e. /θm/ and /mn/ are ill-formed codas and, therefore, require a *repair strategy*. Similar phonotactic patterns in other languages manifest arguments for a syllable level of encoding, such as obstruent devoicing in syllable-final position in German (Wiese, 1996). In Spanish, studies of dialect variation, such as syllable-final /s/ aspiration, nasal velarization, and liquid gliding in the Caribbean dialects (e.g. Guitart, 1997) or gemination in heterosyllabic consonant clusters in Eastern Andalusian (Gerfen, 2001), offer similar support for the construct of the syllable.[5]

Therefore, as we have seen, there are arguments in favor of various levels of phonotactic encoding in the mental lexicon. The specific locus of phonotactics in generative models of phonological grammar is still an active area of current research. However, while the generative approach has contributed a rich body of knowledge to further the field of linguistic inquiry, it is not the only perspective on the representation of language in the mind. The categorical nature of the lexical representations posited by generativists puts them at odds with proponents of other perspectives in which the lexicon consists not of the theoretic-symbolic representation of words and structures but rather actual memory tokens of real speech (e.g. Pierrehumbert, 1999; Bybee, 2001).

1.2.2 Stochastic approaches – An alternative to categorical accounts

Among linguists, an alternative view of the structure of the mental lexicon is reflected in stochastic theories of phonology. These probabilistic models diverge greatly from the traditional generative approach in that they disagree with the basic division of stored lexical items and transformational rules. As described above, the standard practice in generative models of phonology has been to remove predictable information from phonetically realized linguistic

---

[5] However, see Steriade (1999b); Blevins (2003) for counterarguments to syllable-based accounts.

forms, leaving only idiosyncratic (i.e. unpredictable) material to be stored as part of their lexical representations (Bybee, 2001). Variation in natural speech is viewed as related to variation in performance rather than being informative about the underlying structure (Pierrehumbert, 2001a).[6] In contrast, probabilistic models of phonology posit a cognitive representation of sound structure in which frequencies play a crucial role in the acquisition, perception, production and long-term representation. In exemplar-based models, for example, the categorical representations of generative grammars are left aside in favor of memory tokens of individual speech. These memories are organized in a cognitive map in which remembered tokens of high similarity are closer to each other than less similar tokens (Pierrehumbert, 2001b). For example, in order for children to make the generalization that *–ed* often marks the past tense in English, they must first learn a large quantity of verbs exhibiting this form. Once they have acquired the pattern, these words are not necessarily erased from memory. Similarly, new words formed this way may also be recorded in memory (Bybee, 2001).

As multiple tokens of each word are believed to be stored in the lexicon, variability in frequency is able to account for large amounts of empirically observed variation that had been previously ignored by generative linguistic

---

[6] Optimality Theory (OT) takes one step forward towards allowing more information to be stored at the lexical level. While still working within a generative framework, OT does not posit any constraints on input, i.e. the stored form. Rather, surface forms are generated via markedness constraints on the output and/or faithfulness constraints that compare outputs to the input. Therefore, input forms in OT are not the same as underlying representations in traditional generative grammar, and they allow for the possibility of more variation among the items stored in the lexicon.

accounts. In particular, stochastic approaches to phonology are capable of accounting for crosslinguistic distinctions in phonetic space and gradient constraints in phonotactic acceptability patterns. The symbolic nature of generative grammars is not suited for such analyses, as the basic structure of the grammars is categorical. In traditional models, a linguistic property either exists in a language, or it does not. A rule is either applied fully, or not at all. Emergentist approaches, on the other hand, call for a bottom-up approach. As the posited organization of the lexicon is different, statistical probabilities, based on varying similarities among stored items, and varied strength of activation, based on frequency, are able to account for variation of a non-categorical nature.

Bradlow's (1995) comparative study of vowels in English and Spanish found that the vowel spaces for each language differ systematically in the location of their vowel categories in the acoustic space as defined by F1 and F2. Specifically, the comparison of the formant structure of the /i, e, o, u/ vowels in both languages showed that the English vowels are all articulated with a more fronted tongue position (i.e. higher F2) than the corresponding Spanish vowels. This "base-of-articulation" property is a language-specific quality that serves to differentiate between two languages that may share certain phonemic categories. Traditional generative phonology considers the idea of a language-specific articulatory setting outside the area of interest of theoretical linguistics (Bradlow, 1995). For example, in *The Sound Pattern of English,* Chomsky and Halle (1968) consider this aspect of speech as extragrammatical, and thus as part of the performance aspect of language, rather than part of the grammatically determined competence aspect. Yet this is shared knowledge among native speakers that must be represented at some level in the processing of language. Emergentist approaches lend themselves to such cross-linguistic variation given that prototypical formant positioning may be deduced through a base of multiple stored linguistic tokens. Traditional approaches are not capable of such description.

In another crosslinguistic comparison of vowel quality, Laeufer (1992) analyzed vowel duration in recordings of various tokens of English and French monosyllabic words in which the vowels were followed by either voiced or voiceless obstruents. Her findings indicate that, while both languages evidence vowel lengthening, the difference between English and French vowels is not as simple as the presence or absence of a vowel-lengthening rule, given that there is a significant amount of variability in the voicing effect in both languages.

Studies of voice onset time (VOT) exhibit further examples of phonetic variation that is not well captured by categorical accounts. VOTs are the phonetic realization of the phonological concept of voicing. A consonant in any given language may be classified as either voiced or voiceless. The phonological distinction for voicing in French is that between prevoiced stops and short-lag stops. In English the critical boundary is between long-lag stops and short-lag stops. That is to say, what is phonologically considered a “voiced stop” in English may surface voiced with prevoicing but may also appear as voiceless with no aspiration. In a comparative study of VOTs in European French, Canadian French, and Canadian English, Caramazza & Yeni-Komshian (1974) found that

European French and Canadian English correspond to the pattern generally reported for VOT patterns in the two languages. However, the distribution for Canadian French matched neither the European French nor the Canadian English patterns. The VOT values for Canadian French /p/ were comparatively similar in range to European French /p/, but the Canadian French /b/ had more tokens with zero to small positive VOT values and fewer examples of large negative VOT values than those reported for European French /b/. These data show that the VOT range for Canadian French bilabial stops is confined to a narrower range within the total ranges found in European French or Canadian English.

In another comparison of French and English, Flege & Hillenbrand (1986) found that French and English speakers rely on different cues to distinguish post-vocalic /s/ and /z/. Production and perception data were collected on words such as *peas* and *peace*. When the final consonant is voiced, e.g. *peas*, the preceding vowel is longer and the fricative is shorter. While both French and English display the same type of durational differences, their results show that the length of the fricative itself proves to be the stronger cue to distinguishing the word pairs for native speakers of French. English speakers, however, rely more on vowel duration.

The results of these studies and others which examine phonetic variation crosslinguistically make it quite difficult to posit phonetic universals as argued for in generative models. Stated far more strongly in her response to Chomsky & Lasnik's (1995) argument that the PF (Phonetic Form in Chomsky's Minimalist

Program) level of representation is characterized as symbolic, universal, and supporting a uniform interface to the sensorimotor system, Pierrehumbert (1999) contends that

> "Chomsky & Lasnik is not a scientific possibility ... Not only do some phonological entities fail to meet the conditions they lay out, there is no known case of a phonological entity which does meet these conditions. To explain the extremely detailed but extremely systematic patterns which characterize the native phonetics of any language, it is necessary to posit learning mechanisms which can acquire quantitative distributions of phonetic outcomes ... [which] acquire patterns by generalizing statistically over many examples" (p. 114).

Therefore, it is clear from studies such as these that phonetics, fine-grained and gradient in nature, forms an integral part of the linguistic competence. It is demonstrably language specific. The question of possible and impossible phoneme sequences, i.e. phonotactics, carries us from the phonetic domain into phonology. While a traditional understanding of *phonology* implies categorical representations, in the following section I discuss empirical studies that yield important results for a gradient view of phonological modeling as well.

1.2.3 Gradient constraints

Further support for a probabilistic account of the lexicon emerges in studies that have identified gradient constraints in phonotactics. For instance, in

Arabic, verbal roots consist of a set of two to four consonants, with the canonical root consisting of three consonants. Vowels are inserted between the consonants to make word forms. The root /k t b/, for example, forms words dealing with writing, such as كتب, /kataba/ “he wrote”, or كتاب, /kita:b/ “book”. The phonology of Arabic places a co-occurrence restriction on consonantal roots so that there are no roots that repeat the same consonant in the first and second position (e.g. **dadam*) (Frisch et al., 2004). This limitation on possible combinations of consonant forms is accounted for in generative theory through the application of the Obligatory Contour Principle (OCP), a constraint of Universal Grammar, which disallows adjacent similar elements in the consonant root tier (Leben, 1973). According to this approach, the consonants are grouped into natural classes and co-occurrence of segments within a class produces an illicit form. In a study that examines well-formedness judgments of novel verb forms in Arabic, Frisch & Swayed (2001) found that nonce verbs containing consonant pairs that are illicit forms in that they would violate the OCP were judged less word-like than nonce verbs containing consonant pairs that are accidental gaps of equally low frequency. Importantly, this result illustrates the psychological reality of the OCP. However, the judgments were not categorical; the subjects rated the novel forms on a scale of 1-7. The judgments varied on a continuum related to the similarity of the consonants present in the stimuli. They provide evidence that native Arabic speakers have implicitly learned a gradient constraint that reflects a similarity-based consonant co-occurrence restriction (Frisch et al., 2004). This study illustrates one example of how speakers of a language may apply general knowledge of lexical statistics to form judgments on what is possible and impossible in their native language, and more importantly to what degree they would accept novel forms.

Treiman et al. (2000) examined speaker intuitions on possible rime constructions. Specifically, they investigated sensitivity to probabilistic VC constraints in quadruplets of nonsense syllables. Each quadruplet contained two syllables with highly frequent VC combinations and two less frequent syllable forms. The segmental make-up of the quadruplets was controlled in that the high and low syllables contained the same phonemes (e.g. /rup/, /nɝk/, /uk/, /ɝp/). Results showed a reliable difference between High and Low stimuli suggesting that adults consider nonwords with common rimes to be more word-like than nonwords with less common rimes.

In a subsequent experiment the authors submitted participants to a blending task in which the subjects combined two stimuli to form a single item (e.g. /pab/ and /gæf/ → /pæf/, a C/VC blend, or /paf/, a CV/C blend). Results indicate a strong preference for C/VC blends, as is expected for English speakers (Treiman, 1983, 1986). However, responses that broke up the rime (CV/C blends) were more common for syllables with less frequent rimes than syllables with more frequent rimes. Also, more speech production errors occurred on syllables with less common rimes than syllables with more common rimes.

This suggests that adults are sensitive to phonotactic patterns in the lexicon. They implicitly know that some rimes are more probable than other rimes (Treiman et al., 2000). The results do not reflect frequencies of the phonemes themselves, as the high and low frequency syllables contained the same phonemic segments. Native speaker intuitions reflect a detailed understanding of the statistical patterns in the language. These findings contrast with the all-or-nothing rules of traditional linguistic accounts.

In a similar study of internal nasal-obstruent clusters, Hay et al. (2004) found further support for the hypothesis that speakers of a language track statistics across the lexicon. In English some nasal-obstruent clusters are extremely frequent (e.g. /nt/), whereas other are unattested (e.g. /mθ/), and others fall in between the two extremes (e.g. /nf/) (Hay et al., 2004). In their experiment, subjects heard bisyllabic nonce words that contained word-internal nasal-obstruent clusters. They were instructed to rate the novel items as possible additions to the English vocabulary and to transcribe what they had heard in ordinary spelling. As in Frisch et al., 2004, subjects' well-formedness judgments varied on a continuum that mirrors the probability of the nasal-obstruent clusters in the lexicon. Again, the OCP appears to have a cognitive counterpart. However, native speakers' intuitions of permitted combinations are gradient in nature.

These studies and others like them emerge from the growing school of thought in the linguistic literature that the lexicon consists of a very different structure than previously postulated by traditional symbolic theories. The lexicon consists not of purely idiosyncratic material. Multiple speech events of the same symbolic word or phrase may be stored in the mind as multiple tokens. Likewise, linguistic items are not stored in a long, unstructured list. Rather, the regularities and similarities that may be observed in linguistic items are used to structure storage (Bybee, 2001). One way of investigating the cognitive representation of phonotactic knowledge within this framework is through the examination of native intuitions of phonotactic gaps, as we have seen in the research previously discussed. The work elaborated in this dissertation is couched in a probabilistic and emergentist view of the mental lexicon. The following section discusses a specific gap in Spanish: the interaction of diphthongs and stress placement.

1.3 Phonotactic Gaps and Stress in Spanish

How to account for the Spanish stress system is perhaps the research question which has received the most attention in studies of the Spanish phonological system. Of particular interest has been the issue of weight sensitivity. Researchers disagree on the specific interaction of syllable structure and stress placement. This section reviews common observations of Spanish stress behavior, the arguments in favor of and against syllable weight sensitivity in Spanish, nongenerative accounts of Spanish syllable structure, and the special case of the Spanish diphthongs.

1.3.1 The three-syllable window

It is well documented in the literature on Spanish phonology that there is a three-syllable window for stress assignment (Hualde, 2005; Harris, 1995, 1992, 1983; Roca, 1991; Contreras, 1977; Hooper & Terrell, 1976). All nonverbals in Spanish carry primary stress on one of the last three syllables of the word.[7] This gives way to a set of three possible stress patterns in Spanish:

Penultimate stress
a. camisa
árbol

Antepenultimate stress
c. pájaro

Final stress
b. sofá
animal

The previous examples of possible stress patterns are listed in the order of frequency. Núñez Cedeño and Morales-Front (1999) describe that approximately 64% of all words in Spanish carry penultimate stress. It is considered the unmarked stress pattern in the language. Final stress is the second most frequent form (28%), and antepenultimate stress is the least common pattern, being found in only 8% of the lexicon. The low frequency of antepenultimate stress is due to phonotactic constraints which restrict its occurrence. In particular, the composition of the penultimate and final syllables may reduce the three-syllable window to a two-syllable window.

One such pattern is closed syllables. All closed penults will block antepenultimate stress, as seen below.

[7] Certain verbal forms do allow preantepenultimate stress. Gerunds (e.g. *dándomelo*) and imperatives (*pásemelo*) frequently exhibit this stress pattern with enclitics. Nonverbals, however, uniformly fall within the three-syllable window.

Closed penults

a. Consonant in coda position
   i. hablante   ii. *háblante

b. Glide in coda position (falling diphthong)
   i. minotauro ii. *minótauro

Above, starred words exemplify impossible renderings of stress due to the presence of a filled coda or a diphthong in the penultimate syllable. The same process applies to final syllables.

Closed final syllables

a. Consonant in coda position
   i. animal   ii. *ánimal
   iii. Aníbal   iv. *Ánibal

b. Glide in coda position
   i. Paraguay   ii. *Páraguay

The examples listed here and above appear to indicate a case of weight sensitivity in Spanish, insofar as the presence of a heavy syllable in either of the two final syllables of Spanish words reduces the dimensions of the window in which Spanish stress assignment can occur from three to two syllables. In the following section, I discuss in more detail the way in which phonological theory has formally characterized the notion of syllable weight.

1.3.2 Weight sensitivity – Generative approaches

The generative theory of phonological weight, first elaborated by Hyman (1985) within X-theory and further developed by Hayes (1989) employing a moraic approach, distinguishes between syllables in terms of the number of weight units (e.g. moras) they contain. Syllables are classified as either *light* or *heavy*. Hyman (1985) describes that languages that recognize a weight distinction among syllables follow one of two patterns. The first treats a syllable whose rime consists of a short vowel as light and those syllables with rimes consisting of either a long vowel or a final consonant (or more) as heavy. Displayed in moraic representation, this first language type distinguishes syllables as such:

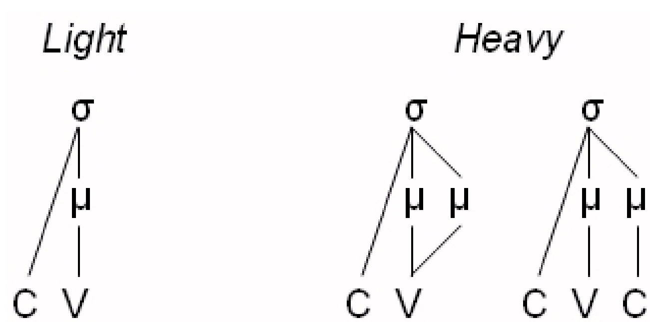

Figure 3: Light and heavy syllables – Type I

The second pattern treats a syllable whose rime has a short vowel as light and a syllable whose rime has a long vowel as heavy, as seen below.[8]

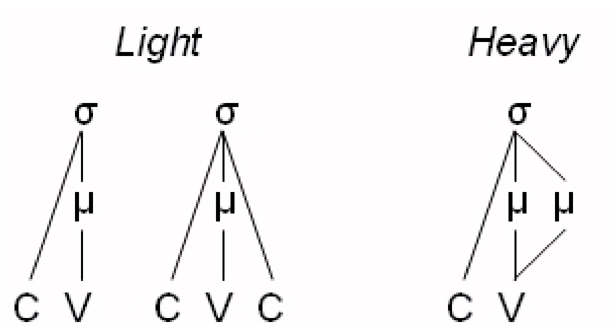

Figure 4: Light and heavy syllables – Type II

Moraic theory serves well to account for syllable weight in that light syllables contain one mora (μ), whereas heavy syllables consist of two. The above typology characterizes the CV syllable as invariably monomoraic and the CVV syllable as invariably bimoraic (Piggott, 1995). The difference between the two patterns is the classification of the CVC syllable. The variability is attributed to a language-specific choice. Hayes (1989) describes that in some languages certain coda consonants are given a mora when they are adjoined to the syllable, phenomenon referred to as *weight by position*. For their part, prevocalic consonants must be parsed as non-moraic onset elements, and never receive weight by position.

Well-known cases of languages in which coda consonants render syllables heavy are most dialects of Arabic, to some extent English, and Latin (Hyman, 1985). In Latin, from which Modern Spanish descends, classifying CVC syllables as "heavy" contributes to accounts of multiple rules and constraints, such as iambic shortening, metrics, and most applicable to this work, stress (Hayes, 1989). Stress in Latin is an example of a generally straightforward trochaic system. Stress is computed from the right edge of the word and falls on heavy penults, or on antepenultimate syllables in the case of a light penult (Mester, 1994). The final syllable in Latin is extrametrical and never carries stress. Mester (1994) illustrates Latin stress placement as follows.

---

[8] See Gordon (2004) and sources cited therein for further discussion of syllable weight typology.

sá pi ēns ‘wise’, nom. sg.
[σ̆ σ̆]<σ>

sa pi én tēs ‘wise’, nom. pl.
[σ̄]<σ>

As depicted above, Latin stress is quantity sensitive. Heavy penults are always stressed. Light penults are always skipped, rendering a proparoxytone pattern.

When we consider modern Spanish, the literature on quantity sensitivity is not conclusive (Bárkányi, 2002; Lipski, 1997; Harris, 1992; Roca, 1991; Den Os & Kager, 1986; Otero, 1986). The question remains whether quantity sensitivity is still active in synchronic Spanish phonology. The majority of the generative work that takes a stance on the matter holds that Spanish is a quantity sensitive language. Examples such as those listed above in which the three-syllable window is reduced to two syllables are listed as demonstrative of the importance of syllable weight to Spanish stress assignment. However, not all researchers are in agreement. Roca (cf. Roca 1997, 1991, 1990, 1988) has written extensively against quantity sensitive approaches to Spanish stress by providing various examples in which the two-syllable window constraints are violated. Native toponyms such as *Frómista* and borrowings such as *Mánchester* and *Wáshington* demonstrate how heavy final and penultimate syllables do not always impede proparoxytone stress. Similarly, Roca argues that simply accenting heavy syllables does not account for the distribution found in Spanish. Specifically, all word-final closed syllables in words with no desinence would be expected to carry the stress, as final syllables are not automatically extrametrical in Spanish, as they were in Latin (Roca, 1991). There are common words in Spanish that end in consonants but carry antepenultimate stress, e.g. *régimen*.

To address these issues, Roca (1991) promotes a reanalysis of the stress bearers in Spanish. Comparing which constraints on syllable structure are violable and which are inviolable, he argues that glides must arise from underlyingly heterosyllabic vocoids which are the possible stress bearers in the language. To support this claim, he lists the proper name *Marcario*. **Márcario* [már.ka.rjo] is clearly not a possible form in the language due to a rising diphthong in the final syllable. While the surface form **Már.ca.rio* consists of only three syllables, and therefore should be a possible word shape (it does not violate the three-syllable window), Roca argues that the motivation for the constraint on stress placement to the final two syllables when a diphthong is present arises from the presence of four underlying vocoids. Specifically, the underlying representation of **Marcario* is /mar.ka.ri.o/, in which placement of stress on *mar*, the first syllable, is outside the three possible stress bearers from the right edge of the word.

In contrast with Roca, Harris (1992) argues in favor of quantity sensitivity in Spanish. In response to Roca's citation of words that violate certain aspects of the two-syllable window (those that derive from consonants in coda position), Harris argues that loan words that maintain their original stress patterns tell us nothing about the rules internalized by the borrowers. He claims that such examples only prove that mimicry of foreign stress patterns is possible. He further stresses that it is the spontaneous pronunciations in which these words appear with final stress that require explanation. Harris claims that this shift to oxytone stress occurs to bring inadmissible stress patterns into conformity with native rules.

Harris (1992) also cites Núñez Cedeño (1986) who finds that in regions of the Dominican Republic where deletion of [s] in coda position is the norm, hypercorrection in which [s] is inserted into words where it never existed, referred to as *hablar fisno,* never produces forms in which the [s] has been inserted into the penultimate syllable of a proparoxytone. For example, the word *hipopótamo* may undergo hypercorrection to produce *hispopótamo, hipospótamo,* or *hipopóstamo*, but never **hipopótasmo*. Harris (1992) argues that the latter example is exactly the pattern in which stress falls on the syllable to the left of a branching rhyme.[9] It is unclear why speakers would exclude only the latter form if the phonological grammar does not prohibit it.

In addition, Harris counters Roca's argument in favor of underlying heterosyllabic vocoids as the source of surface-level glides. He argues that rising diphthongs cannot be underlying heterosyllabic. Words such as *náufrago* and *farmacéutico* show that diphthongs must exist in general before stress is assigned; otherwise the stressed vocoids in these words would fall outside the inviolable three-syllable window (Harris, 1992).[10] Harris also counters Roca's

---

[9] Rather than make reference to moraic structure, Harris prefers to identify what others consider heavy syllables as those which contain branching rhymes in an arboreal representation of syllable structure.

[10] Roca (1991) accounts for proparoxytones with antepenultimate falling diphthongs by claiming that the stress is marked on the antepenultimate underlying vocoid, in the case of *náufrago* on

account by comparing nonverbals like *caricia* with similar verb forms such as *acariciar*. Harris contends that the high vocoid in the final CGV sequence of words such as *ca.ri.cia* must be a glide in the input to the stress rules, because polysyllabic present tense verb forms are stressed without exception on the penult, as in *a.ca.rí.cia* (Harris, 1992). As the root of both the noun and verb derivations must be /karisj-/, Harris establishes that Roca is left with no explanation for the nonexistence and alleged illformedness of nouns and adjectives like **cáricia*, whose stress clearly falls within the three-syllable window.[11]

As can be seen from the above review of two of the most well-known approaches to characterizing the placement of stress in Spanish, it is still unclear in the generative literature if quantity sensitivity remains active in Spanish. What is not in doubt, however, is the reality of the three-syllable window. The debate revolves around how to best characterize theoretically the enforcement of the Spanish stress window. In the next section, I review approaches to Spanish stress that do not take a traditional generative approach, including experimental studies working under computational models and other probabilistic approaches to understanding Spanish stress placement.

---

the *u*, and that when the *a* and *u* become tautosyllabic through Syllable Merger (i.e. the *u* loses syllabicity), the stress is shifted to the *a*.

[11] Here, I have transcribed the root according to standard Latin American pronunciation. Northern dialects of Spain may prefer /kariθj-/.

1.3.3 Nongenerative accounts of Spanish stress assignment

One of the earlier studies to compare generative and probabilistic perspectives in Spanish stress theory is Aske (1990). Aske examines how speakers react to new words presented to them in a way that forces them to choose appropriate stress patterns. Specifically, Aske tests two perspectives: one rising from the generative stance, which he refers to as the disembodied rule view, and another which follows a probabilistic approach, i.e. a patterns-in-the-lexicon view. Subjects were presented with nonce forms written in capital letters and asked to indicate where they would stress the word.[12] Aske developed his stimuli to interact with possible subregularities in the Spanish lexicon. First, he includes items which end in *–en*. In Spanish, nonverbals that end in *–n* are overwhelmingly oxytonic (i.e. stressed on the final syllable), with the exception of forms ending in *–en*, where approximately 62% are paroxytonic (Aske, 1990). If readers follow generative rules, Aske predicts they will stress any new word ending in a consonant (e.g. *–n*) on the final syllable. However, if they are sensitive to patterns in the lexicon, they will analogize the new forms to words with which they are familiar and assign oxytonic stress to words ending in *–an, -in, -on* and *–un*, but primarily paroxytonic stress to the *–en* stimuli.

Another subregularity in the Spanish lexicon is that of words ending in *–ico/a*. The vast majority of vowel-final words are stressed penultimately in

---

[12] This tasks plays on the convenient fact that Spanish orthography does not mark stress diacritically for capital letters. Therefore, a fully capitalized word such as CORUMEN could be read as being stressed on any of the three syllables.

Spanish. However, words with this particular ending typically are proparoxytones, such as *básico, histórico,* or *científico*. Therefore, once again, readers will be placed in a position where they must choose between generative rules for novel forms in which paroxytone stress will prevail, or analogy which will render a proparoxytonic form.

Results show that speakers treat words with *–en* differently from words ending in the other possible *–n* patterns. Aske asserts that this result is incompatible with the hypothesis that speakers make an abstract generalization about Spanish stress, since the subregularities observed do not figure in the formulation of generative rules. Speakers also treat words ending in *–ico/a* very differently from other words ending in a vowel. As predicted, they stressed the words as proparoxytones rather than paroxytones, despite the fact that these words cannot be analyzed as necessarily consisting of a stem plus a suffix (Aske, 1990). Aske thus argues that speakers search the lexicon directly for a suitable pattern when confronted with novel forms. His results do not evidence a case of absolute, abstract generalizations about stress patterns.

Investigating a computational model of Spanish stress, Eddington (2004, 2000) has also tested Spanish stress patterns through statistical patterning. Following Skousen's Analogical Modeling of Language (Skousen, 1989), Eddington submitted the 4,970 most frequent words in the Alameda and Cuetos frequency dictionary (Alameda & Cuetos, 1995) to a test of stress assignment. Building on Aske's work, Eddington examined the computational model's ability to account for the same subregularities and found that the Analogical Modeling of

Language program is able to correctly assign the stress on about 94% of the most frequent Spanish words. Additionally, words that were incorrectly assigned stress by the model are generally those that traditional analyses have treated as exceptional as well (Eddington, 2000). In the *–ico/a* subregularity, the model was able to assign stress accurately to 99 out of 107 words, despite the marked status of the antepenultimate stress pattern. The model also accounted for 83.3% of the *–en* items correctly. This deviates from Aske's finding of 96.8%. Eddington offers the possible explanation for the difference in that the computational model also takes into consideration verbal forms, which would increase the probability of paroxytone stress, whereas Aske's contexts pushed only noun or adjective interpretations. However, when the same stimuli were run against only the nonverbal items in the database, results still assigned penultimate stress to certain items. Therefore, it appears that the model captures native intuitions qualitatively but not quantitatively (Eddington, 2000). While computational models do not imply that the human mind works in the same way, the fact that statistical programs are capable of tracking patterns across a database similar to a native speaker's lexicon provides empirical support to the idea that humans might do the same.[13]

In a different study, Face (2000) examines the role syllable weight plays in the perception of stress by native speakers of Spanish. By means of acoustic manipulation, bisyllabic and trisyllabic nonce words were constructed in which

---

[13] See also Eddington (2004) for further analogical simulations of Spanish stress assignment.

no syllable was more acoustically prominent than any other syllables in the word. Subjects listened to the stimuli and were asked to mark which syllable was stressed. As there were no phonetic cues to pitch or intensity prominence, listener judgments were predicted to follow the unmarked patterns for Spanish stress. Results reveal that speakers did make use of this knowledge in perceiving stress, as there was a strong tendency to perceive penultimate stress when the word ended in a vowel and final stress when the word ended in a consonant. This is the expected pattern to follow if syllable weight plays a special role in stress assignment. That is, a heavy final syllable (ending in a consonant) attracts stress, and a light final syllable (ending in a vowel) is passed over to stress the penult.

Face (2000) also included stimuli that did have acoustically prominent syllables. Subjects correctly identified the stressed syllable 99.5% of the time when the prominent syllable corresponded with the unmarked stress patterns, but only in 82% of the trials where the prominent syllable did not correspond with unmarked stress. Interestingly, the majority of the errors made in the latter condition were instances in which the perceived stress was placed in the unmarked position. These findings in support of syllable weight corroborate studies such as Eddington (2004) that recognize syllable weight as an important component in the Spanish accentual system.[14]

Face (2005) presented the results of the same task used in Face (2000) but with a different subject pool. He found that second language (L2) learners of

---

[14] See also Face (2004) a follow-up study to his 2000 paper.

Spanish are also sensitive to the unmarked stress patterns of the language. Face identifies the unmarked stress pattern in Spanish as final stress for words ending in a consonant and penultimate stress for words ending in a vowel. Specifically, having tested participants at three levels of Spanish language instruction, he found that as learners progress in their Spanish studies so too does their ability to perceive the unmarked stress patterns. However, the data reveal a much higher rate of perceived penultimate stress. This is attributed to the fact that penultimate stress is the default stress pattern regardless of syllable weight factors. Therefore, students acquire the default stress pattern, and later refine their knowledge by acquiring what Face calls the phonologically unmarked stress patterns that are sensitive to the segmental composition of the final syllables of words. Second language learners, like native speakers, thus can be shown to cue in on highly specific phonotactic patterns in their developing L2 lexicons.

Bárkányi (2002) presents subjects with trisyllabic nonce words with varying syllabic structures in a context forcing a nominal interpretation.[15] In an offline written task, participants indicated where stress would fall if the experimental item were a Spanish word. The author argues against quantity sensitivity in Spanish given that not every subject assigned final stress to every nonce word ending in a closed syllable, nor did they all assign penultimate stress to every word ending in a vowel. She states rather that quantity sensitivity is not operative in Spanish anymore, and judgments about the place of stress are not

---

[15] The specific syllable structures were: [CV.CV.CV], [CV.CV.CVC], [CV.CVC.CV], and [CVC.CVC.CVC].

based on strictly defined rules but rely heavily on the lexicon and the analogical influence of existing lexical items (Bárkányi, 2002). Arguably, however, Bárkányi's data speak more to the theoretical discussion of categorical rules versus analogy across the lexicon than specifically to the issue of syllable weight, as a lack of 100% agreement on stress placement offers evidence against categorical generative rule application in favor of analogy across a lexicon in which variance does exist. Her data, in fact, do represent the expected patterns for quantity sensitivity in Spanish. Specifically, it appears from her results that heavy final syllables were most frequently stressed, light and heavy penults were most often stressed when the final was light, and antepenults were rarely stressed unless the penult and ultima were light. The variance cited demonstrates the analogical approach's ability to take into consideration exceptional forms in the lexicon, such as borrowings. As Bárkányi herself admits, there actually are recent loans with a closed penult and stress on the antepenult that belong to what she refers to as a [+borrowing] subgroup, and any word that does not have a proper lexical entry might be interpreted as a word belonging to this group. Thus native speakers are likely to stress proparoxytone stress to 'funny-sounding' words irrespective of the syllabic composition of the word.

1.3.4 Diphthongs in Spanish

In broad terms, it is clear that the issue of syllable weight is a matter of disagreement both in the generative and stochastic/experimental avenues of research. A particularly problematic issue in Spanish phonotactics with great import on theories of Spanish syllable weight are the diphthongs. As mentioned above, falling diphthongs tend to pattern with syllable-final consonants in that they shrink the three-syllable window to one of two syllables, and they also attract stress themselves. As noted above, researchers disagree vehemently as to the theoretical underpinnings that motivate such patterning, but descriptive accounts indicate that, although there are exceptions, syllables containing a coda consonant and falling diphthongs tend to be assigned stress from the right edge of the word leftward.

However, the interaction of diphthongs and stress becomes less clear when we examine how rising diphthongs affect stress patterns in Spanish.

Rising diphthongs

Prevocalic glide in penultimate syllable
i. va<u>ria</u>ble ii. *<u>vá</u>riable

b. Prevocalic glide in final syllable
i. fa<u>mi</u>lia ii. *<u>fá</u>milia

These examples illustrate how antepenultimate stress is blocked by the presence of a rising diphthong in either the penultimate or final syllable. The three-syllable window is clearly reduced to two. In this sense, rising diphthongs pattern with falling diphthongs in that they block antepenultimate stress. However, unlike falling diphthongs, rising diphthongs do not necessarily attract stress themselves, as seen in *familia*. In *familia,* we find a rising diphthong in the final syllable, yet stress falls on the penult. If rising diphthongs attracted stress, *familia* would be stressed on the final syllable. This is clearly not the case, given the wealth of Spanish paroxytones with rising diphthongs in the final syllable. A search of the Alameda & Cuetos (1995) database revealed 803 words of this type, such as *sitio, ambigua, especie,* etc.

A possible comparison to be made is between rising diphthongs and complex onsets. Take for example the following words.

Rising diphthongs vs. complex onsets

a. aduana
b. temprana

Here we see two words, one containing a rising diphthong in the penultimate syllable (a.) and the other a complex onset (b.). They share the same final syllable and carry penultimate stress. It appears that they pattern identically. However, if rising diphthongs and complex onsets pattern equally, then the latter should also block antepenultimate stress. There are, however, many words in the Spanish lexicon that permit antepenultimate stress with a complex onset in the penult, as seen below.

Antepenultimate stress with complex onsets

a. biógrafo c. rúbrica
b. lágrima d. república

This observation led Harris (1983) to consider rising diphthongs an example of a branching nucleus in his treatment of Spanish syllable structure.[16] According to his account, antepenultimate stress is impossible with *either* a branching nucleus

[16] See also Hualde (1991) and sources cited therein for corroborative evidence in favor of the nuclear status of prevocalic glides.

or a branching rime (a consonant in syllable-final position fills the coda position, causing branching at the rime level, rather than nuclear level). Therefore, a distinction is made between rising and falling diphthongs and syllables closed by a consonant. Roca (1991) observes that there are multiple violations of the two-syllable window when the final syllable is closed by a consonant, such as *Wáshington, Mánchester, Frómista*, etc. However, he claims there are no exceptions to the two-syllable window when the penult contains either a rising or falling diphthong. It appears that Harris's distinction between branching nuclei and branching rimes accounts for this observation. That is, branching rimes in the penult or ultima may occasionally permit antepenultimate stress, but a branching nucleus never does.

However, neither Harris' nor Roca's theory of syllable structure differentiates between the rising and falling diphthongs themselves. Roca makes no claims about the difference because he does not believe that quantity sensitivity is active in Spanish and therefore needs to make no claims about syllable weight. Harris does argue in favor of quantity sensitivity but puts forth no direct theoretical claims about the differential behavior of the two diphthong types. He simply argues that they both represent branching nuclei and reduce the three-syllable window to two syllables.

A usage-based approach to the diphthong conundrum would argue that the differential behavior in diphthong types is a reflection of the patterns in the lexicon. The extant patterns may be due to historical restrictions on syllable structure in Latin that have been carried forward to Modern Spanish. It is also possible that the synchronic phonological system currently constrains the interaction of rising diphthongs and stress placement. It is the focus of this dissertation to investigate further the cognitive representation of these restrictions. Therefore, in the following section I discuss psycholinguistic models that consider the cognitive processes involved in reading and speaking. A clear understanding of the cognitive mechanisms to produce speech is necessary in order to detail the methodology adapted in this dissertation and as a point of departure for the exploration of the representation of phonology and the encoding of phonotactic constraints in the mind.

1.4 A Behavioral Approach – Psycholinguistic Models of Reading and Speech Production

The goal of this dissertation is to take an experimental approach to examining how phonotactic restrictions on stress placement are represented cognitively by native Spanish speakers. A behavioral approach allows for the collection of online data that probes native speaker knowledge by recording their behavior when presented with carefully developed stimuli. Unlike many of the studies discussed thus far in this review, online tasks elicit responses from participants while minimizing interference from metalinguistic intellectualization. The present work adapts methodologies from psychology to collect behavioral data to test theoretical hypotheses regarding the nature and status of phonotactic constraints empirically.

In this section, I will examine two theories of reading and two theories of speech production in the psycholinguistic literature. I end the review with a discussion of empirical evidence from psycholinguistic studies on Spanish.

1.4.1 Models of reading

Despite the traditional linguistic approach of ignoring the written form for the study of phonological grammar, it is important from a cognitive perspective to understand the processes that readers undertake when reading words aloud. The experimental tasks to be discussed in subsequent chapters of this dissertation consist of reading aloud words presented on a computer screen. It is therefore relevant and necessary to understand psycholinguistic models of reading. When reading, we are presented with a visual stimulus. As literate speakers of our language, we must convert the symbols in the stimulus into meaning and sounds. This process is quite complex, and it is the task of psycholinguistic models to explain the cognitive mechanisms that allow us to process written language.

The current debate in the word naming literature is between localist dual-route models (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001; Coltheart, Curtis, Atkins & Haller, 1993) and distributed connectionist models (Plaut, McClelland, Seidenberg & Patterson, 1996; Seidenberg & McClelland, 1989). Both models can account for most of the data in human word production, but their fundamental explanations of the paths through which orthography is transferred to phonology differ greatly.

1.4.1.1 The dual-route model

Coltheart's dual-route model claims that there are two routes through which activation of a word and its subsequent transfer to speech may occur.[17]

[17] Figure of Coltheart's model adapted from Coltheart et al. (2001).

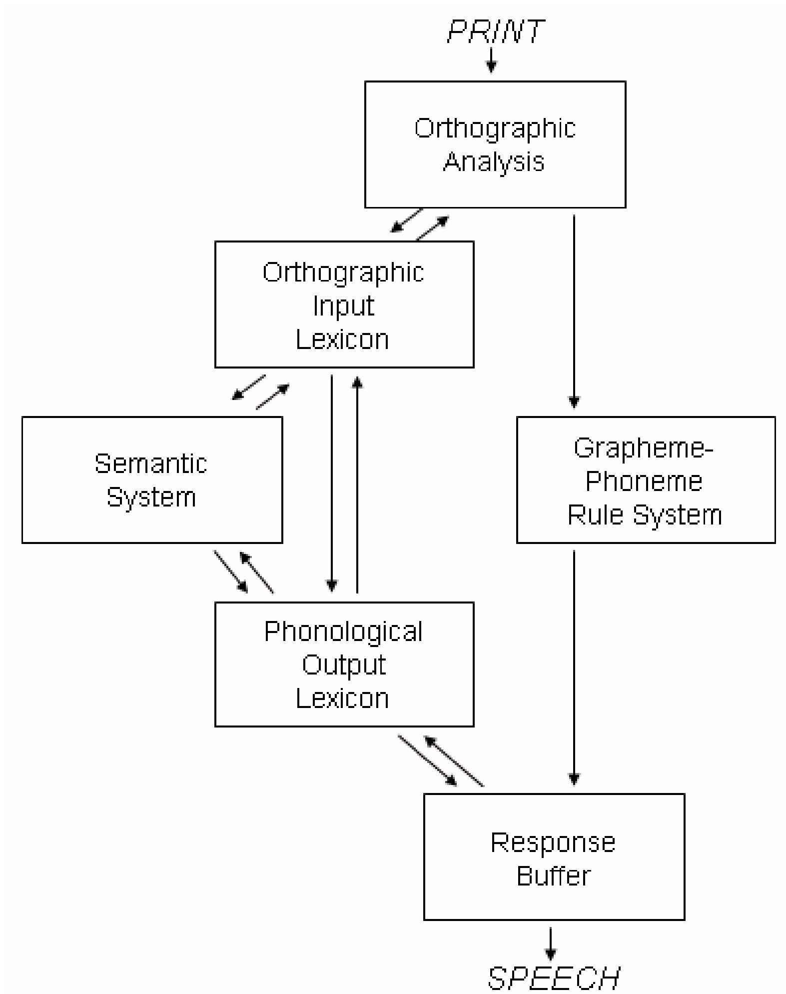

Figure 5: Coltheart's dual-route model of word naming

Both routes consist of three levels of activation. When a word is read, visual features of the print are analyzed and fed to a letter-unit level. At this point the model diverges onto two paths. In the nonlexical route the activation feeds to a grapheme-phoneme rule system. Grapheme-phoneme correspondence (GPC) rules convert the letter stream from left to right into phonemes. The GPC rules are context-sensitive to the position of the graphemes in the word and consider phonotactic and monomorphemic constraints. The activated phonemes are analyzed in a phoneme system and converted into speech.

In the lexical route of Coltheart's model, once the visual feature units have been analyzed at the letter unit level (two levels shared by both routes), the activation is fed onward to a subsequent level in which the word's entry in the orthographic lexicon is activated. This word entry then activates the corresponding word entry in a phonological lexicon. The word's phonemes are subsequently activated in parallel across all phoneme positions allowing for speech (Coltheart et al., 2001). The model also consists of a semantic system that may play a role in word activation between the orthographic input and phonological output lexicons.

Irregular words (i.e. those whose pronunciation cannot be derived from GPC rules) have frequently been cited in the literature as yielding longer naming latencies than regular words (Seidenberg, Waters, Barnes & Tanenhaus, 1984; Coltheart et al., 2001). In Coltheart's model, irregular words rely heavily on the lexical route of activation. Regular words are more quickly activated through the nonlexical route. Given that both routes of the dual-route model are active simultaneously to process print into speech, inhibitory effects occur when both routes meet at their third shared level, the phoneme system, where discrepancies in activation will slow speech production. For example, the irregular word *yacht*, will activate the correct pronunciation /jat/ once its entry is accessed in the lexical route. However, the nonlexical route will activate /jæt͡ʃt/ through GPC rules. Therefore, the two competing forms must be reconciled in the phoneme system before the word can be pronounced, yielding higher reaction times in naming experiments for irregular words.

Rastle & Coltheart (1999) found that the size of the latency incurred by irregular words in word naming declines as a function of the left-to-right position of the grapheme-phoneme miscorrespondence in the word. The nonlexical route of the model activates phonemes sequentially from left to right, as opposed to other models of word naming in which all phonemes are activated in parallel (Coltheart, 2001). Coltheart et al. (2001) view this evidence as support for their model, which they claim poses a particular challenge for the connectionist models of word production.

1.4.1.2 Parallel connectionist models

Parallel connectionist models of word naming (Plaut et al., 1996, Seidenberg & McClelland, 1989) propose a very different view. In such models, readers learn the statistical relationships between spelling and sound, rather than rules (Chateau & Jared, 2003). Orthographic, phonological, and semantic information is represented in terms of distributed patterns of activity over separate groups of simple neuron-like processing units (Plaut et al., 1996). The connections between the units that represent spelling and sound are controlled by weights.[18]

---

[18] Figure of parallel connectionist model adapted from Seidenberg et al. (1996).

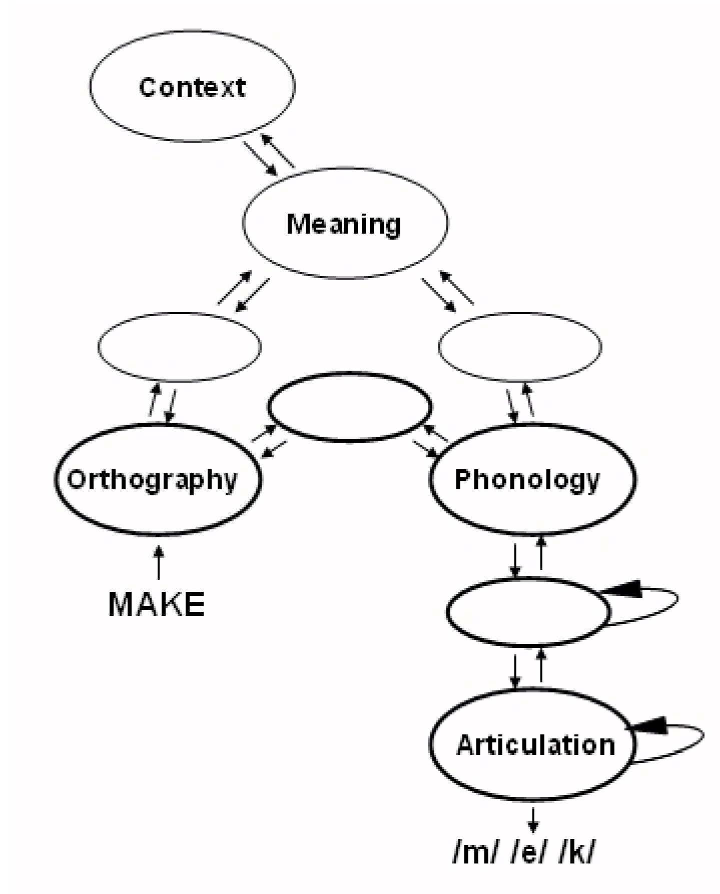

Figure **6**: A parallel connectionist model of word naming

This model differs drastically from the dual-route model in that it calls for no representations of individual words (Jared, 2002). Rather than access word entries in an orthographic lexicon, as in Coltheart's model, Plaut et al.'s model claims that the lexical status of a string of letters or phonemes is not reflected in the structure of the reading system. Rather, words are distinguished from nonwords only by *functional* properties of the system, i.e. the way in which particular orthographic, phonological, and semantic patterns of activity interact. This model also differs from other models of word naming in that it does not assume separate levels of interaction for irregular and nonwords. Conflicts among possible pronunciations of sound strings are resolved by cooperative and competitive interactions based on the correlation between the letter stream being read and all known words and their pronunciations (Plaut et al., 1996). Therefore, the ease of pronouncing a word depends on the relative consistency of the pronunciation of the letter patterns in the word. The less typical the pronunciation in the inconsistent word, the greater the difficulty with which it will be pronounced (Jared, 2002).

The models also differ with regard to regularity and consistency. The distinction between these two terms is important when testing models of word production when reading. Regularity, as mentioned previously, refers to a correspondence between the orthographic and phonological representations of a word. A word is regular if its pronunciation can be directly derived from its spelling. Irregular, or exceptional, words are those whose pronunciations do not reflect a direction correspondence to orthography (such as the *yacht* example mentioned earlier). Consistency, in contrast, refers to the relationship between a word and other words with similar spelling. A word is considered inconsistent if other words that share the same word-body have conflicting pronunciations, such as *waste* and *caste* (Glushko, 1979; Jared & Seidenberg, 1990; Chateau & Jared, 2003). That is, if words share the same orthography representing distinct phonological rimes, they are considered inconsistent. Word neighbors with conflicting pronunciations are referred to as “enemies”, and words with a shared pronunciation are “friends.” Both regularity and consistency cause effects in naming latencies and are controllable variables in naming experiments.

Andrews (1982) found main effects of consistency and regularity in naming latencies for low-frequency words, but only consistency effects for high-frequency words. Therefore, we can assume that these two variables, although they often lead to similar effects (i.e. variation in naming latencies), are not the same. Coltheart's model argues that consistency effects arise due to the activation of words that are orthographically similar to the target in the orthographic lexicon level of the lexical route. Regularity effects, because they refer to the degree of grapheme-phoneme correspondence, are a result of the nonlexical route of the model. Jared (2002) argues that it is possible that GPC rules are not used in naming words and that word-body consistency can account for the naming data (which would reflect the connectionist view that orthographic patterns within the distributed network alone can account for these effects). In order to test this hypothesis Jared ran a series of experiments.

Jared's (2002) experiments manipulated word regularity, consistency and frequency. Frequency was deemed an important variable, as previous studies established different results for low versus high frequency words. The dual-route framework argues that high-frequency words are quickly retrieved from the lexicon before information from the nonlexical route becomes available (Seidenberg, 1985). However, for low-frequency words, because retrieval from the lexical route is slower, phonological and lexical information becomes available at similar times. Therefore, latencies are increased for low-frequency irregular forms, because a conflict arises between the incorrect phonological representation derived through the nonlexical route and the correct pronunciation produced by the lexicon. The connectionist model contends that the weights on the connections between the word's orthographic and phonological patterns are updated frequently enough in high-frequency words to reduce or eliminate the noise in phonological encodings (Jared, 2002). Therefore, the processing of inconsistent orthography-phonology correspondences becomes more precise in high-frequency words.

In her first experiment, Jared (2002) compared the performance of irregular and regular-inconsistent forms with regular-consistent forms controlled for frequency. She found an effect of consistency for low-frequency words with little evidence of any effects of regularity. Therefore, the results suggest that word-body consistency better characterizes the difficulty associated with naming a word than GPC regularity. These results challenge the dual-route model in that consistency is defined in relation to the orthographic body of a word or nonword, a level of representation absent from the dual-route model (i.e. Coltheart's model accounts for the relationship between individual graphemes and phonemes, or entire word entries in the orthographic lexicon. No intermediate level represents orthography body units in his model.).

Jared's (2002) subsequent experiments find that when frequency and neighborhood effects are unconfounded, frequency and regularity interactions are fairly weak (experiment 2), and when high and low frequency words are matched for neighborhood characteristics, frequency appears to have much less impact on consistency and regularity effects than previously assumed (experiment 3). Accordingly, she argues that regularity effects depend on neighborhood characteristics, which supports the view that word-body consistency is a stronger argument for variation in naming latencies. This in turn suggests that readers learn more than individual letter-sound correspondences. They also learn how letters are pronounced in their surrounding context (Jared, 2002). This pattern of statistical learning rather than spelling-sound rules supports the connectionist model of word naming.

1.4.1.3 Pseudohomophone effects

Another area of research in word naming that lends insight into the difference between the two models is the issue of pseudohomophone effects. Pseudohomophones are nonwords that sound like real words (e.g. *brane* or *joak*). The comparison of these forms with nonpseudohomophones (nonwords that do not resemble real words phonologically) provides a way to diagnose the activation of phonological information in reading (Seidenberg, Petersen, MacDonald and Plaut, 1996). Studies conducted by McCann & Besner (1987) and McCann, Besner & Davelaar (1988) found that subjects were faster to name pseudohomophones compared to nonpseudohomophones. These findings were interpreted as suggesting that the pronunciation of a pseudohomophone involves the activation of the base word from which it is derived, causing the facilitative effects in the naming task. They replicated these findings in a lexical decision task. Results showed that pseudohomophones yielded longer latencies than nonpseudohomophones. This was taken to mean, once again, that the base word was activated in the lexicon, this time causing inhibitory effects. That is, activation of the base form facilitates pronouncing the word, but it interferes with concluding that the stimulus is not a word (Seidenberg et al., 1996). This interpretation assumes, then, that the mental lexicon must hold individual lexical entries for words. Otherwise, it would be impossible for the base words resembling the pseudohomophones to interact in the naming and lexical decision tasks. Such reasoning is difficult to accommodate in the connectionist model in which patterns are distributed across a network with no individual word representations.

Taft & Russell (1992) found similar results in a study in which the frequency of base words was also controlled. They found frequency effects in word naming for pseudohomophones, but only when the readers were slow. Interestingly, faster readers showed neither effects of pseudohomophone base-word frequency nor an advantage for pseudohomophones at all. Coltheart et al. (2001) interpret these results as support for the dual-route model. The speed of the nonlexical route regulates the pseudohomophone effect in that there is more time for activation of the base word in the phonological lexicon to rise when the nonlexical route operates slowly. This activation feeds forward to facilitate activation at the phoneme level. Given that activation in the phonological lexicon is affected by word frequency, pseudohomophones whose base-word frequencies are high will enjoy more facilitation than those with lower base-word frequencies (Coltheart et al., 2001).

Seidenberg et al. (1996) offer a different perspective. They maintain that the pseudohomophone effect does not arise due to the fact that pseudohomophones resemble real words and nonpseudohomophones do not, but rather that pseudohomophones are more similar to real words in other respects. The authors claim that creating nonwords that sound like words tends to require using components that occur more often in actual words. Also, nonwords that are more similar to actual words are more likely to be pronounced more easily than less word-like nonwords. This leads to more difficulty in discriminating them from words in a lexical-decision task (if we assume a connectionist perspective in which lexical decisions are made based on statistical patterns of real words). Therefore, Seidenberg et al. (1996) argue that word-specific representations are not required to account for pseudohomophone effects in naming and lexical decision tasks. Effects of lexical frequency (i.e. facilitatory effects in word production due to higher frequency among base words of pseudohomophones) are accounted for in the articulatory domain in connectionist models. Nonword pronunciations are lower frequency patterns than word pronunciations. Therefore, nonwords that more closely resemble the higher-frequency patterns will be named more quickly, given that lexical frequency is represented as being an impact on the knowledge represented in the network by varied weights in the phonological system rather than the knowledge of word-specific representations (Seidenberg et al., 1996).

As we have seen in this section, there are two main competing perspectives on the processing of printed words and converting them to speech. The localist dual-route model (Coltheart et al., 2001) calls for individual word representations which are accessed through a lexical route at the same time as a second nonlexical route converts the graphemes into phonemes based on phonotactic and monomorphemic rules. The parallel distributed model (Seidenberg & McClelland, 1989; Plaut et al., 1996) posits the existence of a neural network in which statistical patterns of orthography, phonology and semantics are interconnected through weighted hidden units.

Although both models are capable of accounting for the majority of the data presented in previous studies, the advocates of each model have found results which challenge the opposing model's views. Coltheart et al. (2001) have argued that positioning effects of grapheme-phoneme asymmetries are better accounted for in their model, due to the serial nature of the nonlexical route. Jared (2002) argued that consistency effects more clearly characterize the differences in naming latencies between consistent and inconsistent words. She asserts that the connectionist model, which allows for statistical learning of patterns at an intermediate level between grapheme-phoneme correspondence and the complete word level, offers a more logical explanation of such results. Both models account for pseudohomophonic effects by focusing on different aspects of the similarity between pseudohomophones and real words (i.e. they either activate base words in the lexicon or they activate more frequent patterns in the distributed network).

While the goal of this research is not to adjudicate between models of reading, it is important to realize that within either model the activation of phonological encoding plays a central role. It is clear that the methodology employed in the previous studies is a viable means of probing phonological knowledge.

1.4.2 Syllable encoding in psycholinguistic models of speech production

One pertinent aspect of word production typically ignored by the previously discussed models is the encoding of syllable structure. As described in sections II and III, the syllable is intrinsic to the linguistic study of the phonotactic gaps of interest to this dissertation. As the work here examines lexical gaps that hinge on the interaction of syllable structure and stress, it is important to examine how psycholinguistic models account for the encoding of such knowledge.

While the previous models of reading focused more closely on the conversion of orthographic forms into phonology, other models focus more specifically on the step-by-step process of speech preparation and production. The most prominent psycholinguistic models of speech production are the Levelt, Roelofs, and Meyer model (1999) and the model proposed by Dell (1986, 1988). Both models structure the lexical into three primary levels. When preparing to speak, the first step in the process is to activate concepts in a conceptual level. Activation from this first level feeds to a second level in which a word (or words) that corresponds to the activated concepts is selected. The "word" in the second level is a symbolic conception of the word to be spoken. In the third level, the phonology associated to the word is encoded. Once this encoding is complete, the word is articulated.

The division of speech production into these three levels is quite intuitive. As speakers of languages, we are able to identify similarities between words. We can classify words into groups based on different characteristics (*banana, apple,* and *orange* are fruits; *dog, cat,* and *mouse* are animals). This is possible because words share conceptual information. When we want to name a particular animal, conceptual nodes associated with that animal allow us to select the appropriate word from the mental lexicon. This marks the boundary between the first and second levels of speech production. Similarly, we are all familiar with tip-of-the-tongue situations in which we *know* the word we want to say, but we are not able to produce it. Studies such as Vigliocco et al. (1997) show that speakers of language which mark nouns with grammatical gender (such as Dutch or Italian) are able to identify the gender of the word they want to say, even if they are not able to produce the word itself. This illustrates the division between the second and third levels. While both models agree on this general treatment of speech production preparation, they differ greatly on the specific representation of linguistic knowledge in the mind, and the interaction between the different levels.

1.4.2.1 Dell's two-step interactive activation model

Dell's model is often referred to as the two-step interactive model. A visual representation of the model can be seen below.[19]

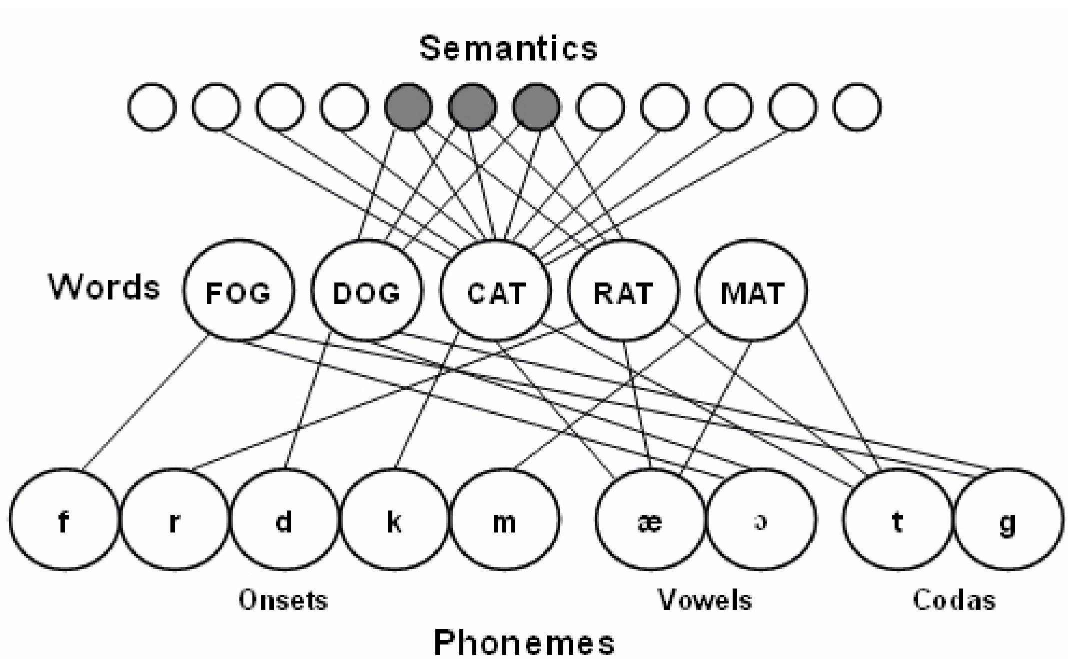

Figure 7: Dell's two-step interactive activation model

The model is referred to as "two-step" because it consists of the previously discussed three levels, with two steps of connections between them. The connections between the levels are bi-directional, meaning that activation can spread in a top-down, or bottom-up direction. This is typical in a connectionist model and is able to account for many types of speech errors such as those discussed below. All connections are facilitatory, and there is free cascading of activation throughout the network. The most highly activated form at each level

[19] Figure adapted from Dell et al. (1997).

is selected. This selection is made by external factors. For example, if a person is presented with a visual object to name, those semantic nodes consisting of concepts that match the object will receive a jolt of activation. This feeds to the word level. The previously selected conceptual nodes feed activation to various words. Here, word selection is also determined externally by a syntactic frame. A jolt of activation is applied to the most activated word, giving that candidate an advantage in phonological encoding. The word's phonological encoding proceeds in a similar fashion in which multiple phonemes receive activation from higher levels and are selected into syllable frame slots. The final stage in Dell's model is the translation of the most highly activated phonological forms into articulatory codes for speech.

1.4.2.2 Levelt, Roelofs and Meyer's serial discrete model

The strongest opponent of Dell's model of speech production is the model proposed by Levelt, Roelofs, and Meyer (1999). This model is similar to Dell's in that it is a non-distributed model in which words are stored as whole linguistic units in the mind. These nodes are selected based on activation from a higher conceptual stratum, and they in turn feed activation to a lower form stratum. A visual representation of Levelt et al.'s model is seen below.[20]

[20] Figure adapted from Levelt et al. (1999)

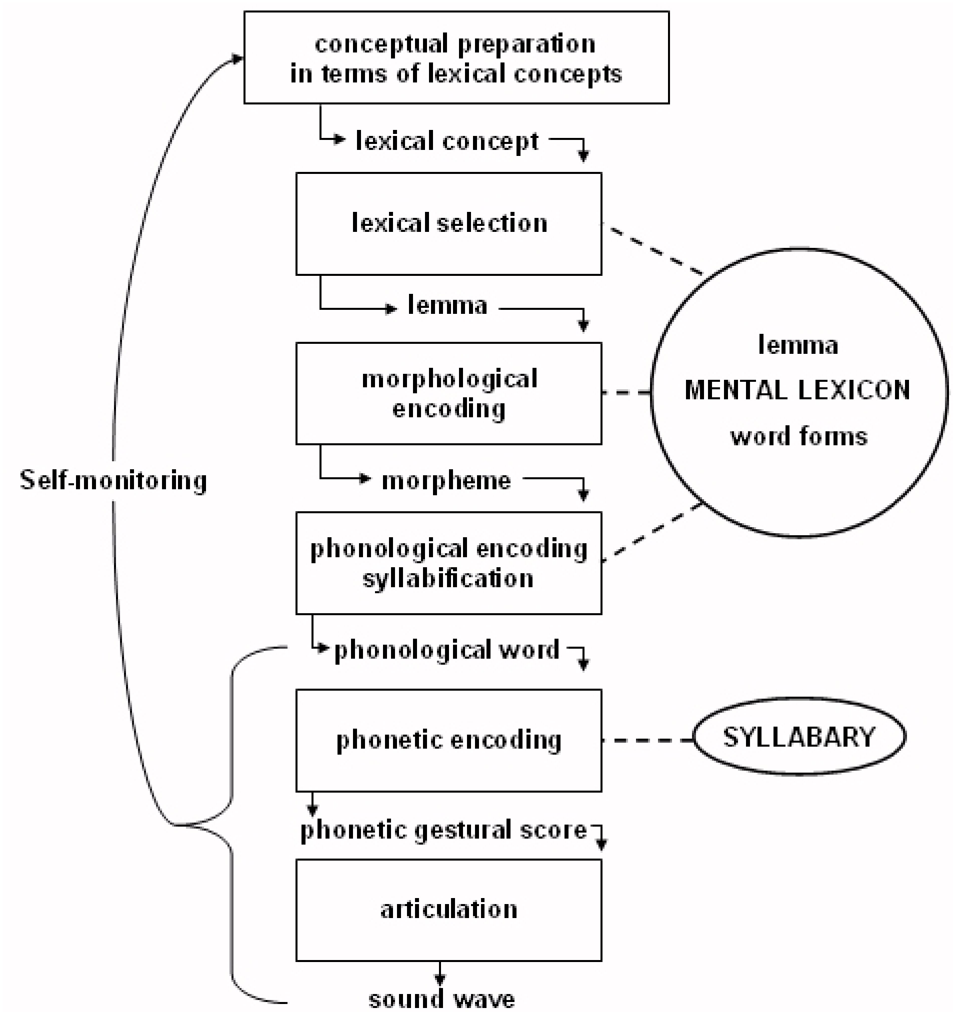

Figure **8**: Levelt, Roelofs and Meyer's serial discrete model

While this model looks very similar to Dell's model, there are crucial differences between the two. In Levelt et al.'s model, there is only feed-forward activation. That is to say, forms activated in lower levels cannot spread activation back up to previous levels. The model is also modular in that only one selection is made at each level (as opposed to the cascading model in which activation spreads freely). Therefore, in this serial model of speech production, lexical concepts in a conceptual stratum lead to a lexical selection level most often referred to as the lemma level. Here, a single word (or *lemma*) is chosen from the mental lexicon.

According to this model words are stored in the lexicon with syntactic and morphophonological information. Syntactic information is stored in the form of diacritic features. For example, verbs in English are stored with features for number, person, tense, aspect, and mood. That is to say, the lemma *escort*, for example, may be phonologically realized as *escort, escorts, escorted,* or *escorting*, depending on the values of its diacritic features (Levelt et al., 1999).

Moving from the lemma stratum to the form stratum, the next step in the preparation of the word is to encode its morphological makeup, its metrical shape, and its segmental makeup. The morphology of the word is selected by choosing the form that corresponds to the appropriate syntactic features (for example, *escorting* among the previously listed options). In the case of *escorting*, two morphemes are identified: *escort* and *ing*. The metrical shape for these two morphemes is (σ σ́) and (σ), respectively. Finally, the phonemes /ə/, /s/, /k/, /ɔ/, /r/, /t/, /i/ and /ŋ/ are selected for activation. The selection processes in the lemma and form strata are illustrated below.[21]

[21] Adapted from Levelt et al. (1999).

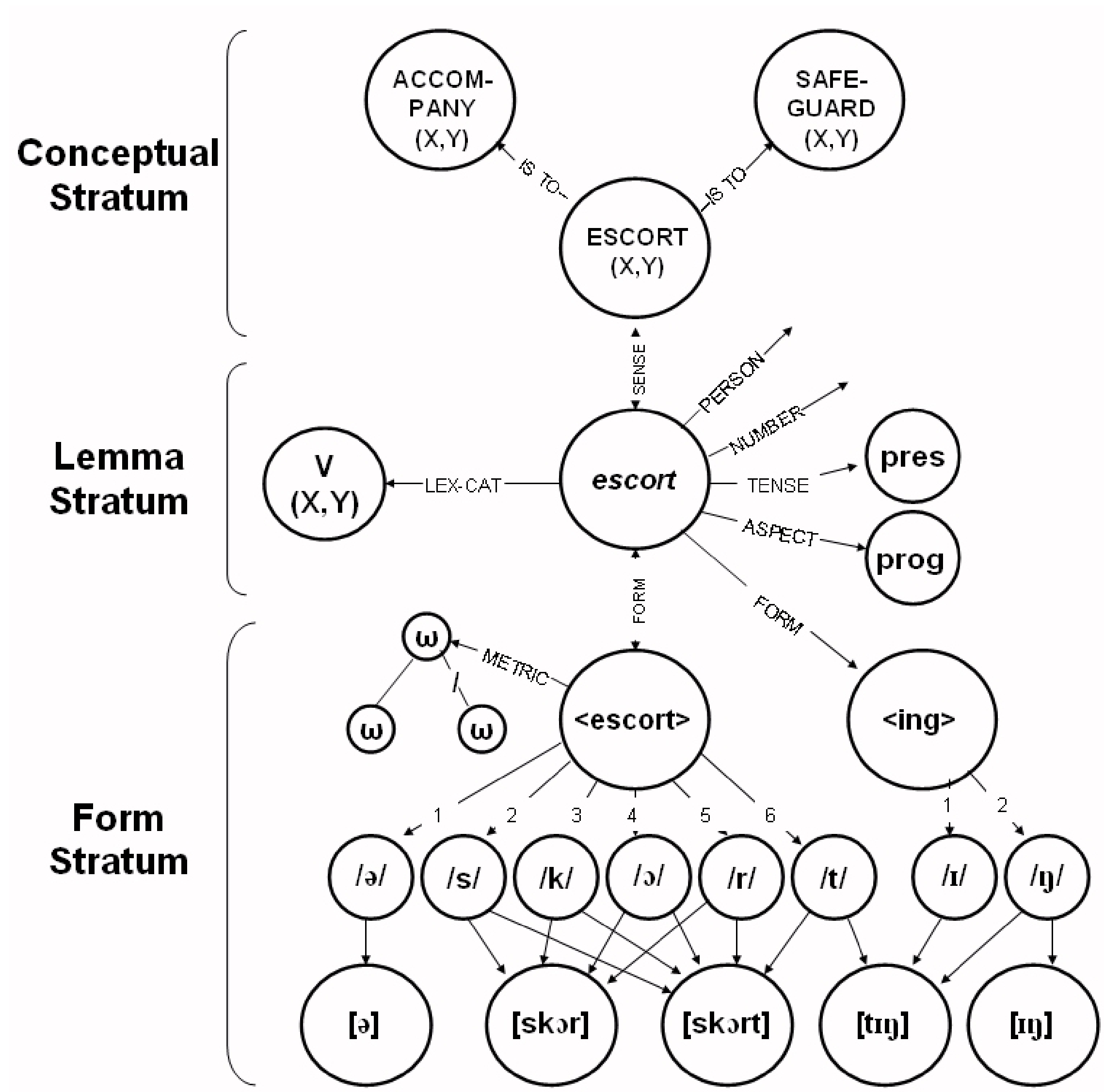

Figure **9**: Selection processes in Levelt, Roelofs, & Meyer's model of speech production

As opposed to Dell's model in which syllable structure is stored in the lexicon, Levelt et al. posit that syllables are stored in a mental syllabary. During the phonetic encoding of the phonological word, speakers access the syllabary to retrieve stored articulatory gestures of highly frequent syllables. Studies from corpus data show that speakers of languages such as English or Dutch do most of their speaking with only approximately 500 syllables despite having over 10,000 possible different syllables in the languages (Levelt, 1999). Schiller et al. (1996), for example, found that 500 syllables account for 84.75% of all syllable tokens in the CELEX database. Therefore, Levelt et al.'s model proposes that these highly frequent articulatory gestures are stored in a syllabary so as not to have to recompute them every time they are required for word production. Hence, a large distinction between the two speech production models is the processing of syllabic information. Dell's model conceives of a lexicon in which words are stored with syllable encoding, whereas Levelt et al.'s model argues that words are only stored with larger-scale metrical information, such as stress placement. In the latter model, syllable structure is computed online during phonetic encoding of the phonological word prior to articulation.

A final component to Levelt et al.'s model is the self-monitoring system. As illustrated in figure 8, the model illustrates our ability to monitor our own speech preparation and production. As the authors write, the person we all listen to most is ourselves. We are able to monitor the phonetic encoding and articulation of our own speech. In the case of a detected error, we are able to re-encode the word to correct the error. Similarly, tasks that distract our ability to monitor phonetic encoding and articulation may lead to variance in naming latencies and/or accuracy.

1.4.2.3 Empirical support for Dell's model of speech production

Two arguments used frequently in the literature to argue in favor of a connectionist framework are the lexical bias effect and the mixed error effect. Early studies that examined the quality of production errors found that when speakers incorrectly named a word, they produced another real word more than would be expected by chance (Baars, Motley, & MacKay, 1975; Dell, 1986; Stemberger, 1985). That is to say, when prompted to pronounce *cat,* a person is more likely to produce *mat* than *gat*. Interactive models such as Dell's account for this lexical bias effect in the following way. Once the word *cat* has been activated in the word level, it will feed activation to its corresponding phonemes in the subsequent level. Given that Dell's model is interactive, the activation produced in the phonemes /k/, /æ/, and /t/ will spread back up to the word level to feed activation to all words that share those phonemes (*hat, fat, mat*, *cap,* etc.). These words will then in turn feed activation to their phonemic counterparts. This will further activate the shared phonemes in *cat* but will also increase competition by activating new phonemes not previously activated (/h/ in the case of *hat,* /m/ in the case of *mat*, etc.). Dell (1986) refers to this flow of activation as positive feedback loops. Of course, nonwords, such as *gat*, have no counterpart at the word level and, therefore, receive no added boost in activation as the real words do. If there is a breakdown in the phonological process, competitors in the loop, such as *mat* or *hat*, are more likely to be articulated than nonwords such as *gat* or *zat*. Connectionist approaches argue that feedback within the system is necessary to account for such an effect.

Proponents of serial models, such as Levelt et al.'s model, respond that the lexical bias effect can also be explained through the self-monitor mechanism. If we assume that we monitor our own speech and are able to catch production errors at the phonetic level of encoding, it is arguable that real words are more likely to slip past the monitor. Also, Levelt argues that bi-directionality needs independent motivation. In his words, "its functionality can hardly be to induce speech errors" (Levelt, 1999). He further argues that one should consider the possibility that bi-directionality within the system very well may be a property of the error itself. That is to say, the errors might occur precisely because interactivity has arisen in an otherwise discrete system (Levelt, 1999). Further support for the self-monitor account is the fact that lexical bias effects are strongly influenced by circumstances such as task demands and conversational settings. The self-monitoring mechanism is also hypothesized to be sensitive to such adjustments. Baars et al. (1975), for example, found no lexical bias effect when all of the stimuli were nonwords (Levelt, 1989).

The mixed error effect refers to the additional observation that semantically similar production errors also tend to be similar phonologically (Rapp & Goldrick, 2000). For example, if a speaker wants to say *cat,* the conceptual nodes associated with it will also activate semantically related words at the word level, e.g. *dog, rat,* or *bird*. Given that Dell's model assumes cascading activation, these forms will also spread activation further to their associated phonemes (*dog* = /d/, /ɔ/, /g/; *rat* = /r/, /æ/, /t/; *bird* = /b/, /ɚ/, /d/). Assuming that the phonemes associated with *dog, rat,* and *bird* receive relatively equal amounts of activation from the word level, *rat* should have an advantage in that it also shares two phonemes with the target word *cat*. Therefore, if a production error were to occur, the probability that the word spoken instead of *cat* will be both semantically and phonologically similar to *cat* is higher than chance (Goldrick, in press).

Levelt et al.'s model does not allow cascading activation between the lemma and form levels. The model is discrete; only the phonemes for the selected lemma receive activation at the form level. Connectionists argue that the assumption of cascading activation is the best way to account for the mixed error bias effect. Levelt and colleagues claim that the self-monitoring mechanism can account for the mixed error effect in the same way that it is able to account for the lexical bias effect. Those errors that are both semantically and phonologically similar to the target form are less likely to be detected by the monitor than less-similar forms (Roelofs, 2004). Therefore both models offer differing accounts of the mechanisms resulting in speech errors such as lexical bias and semantic-phonological mixed effects.

1.4.2.4 Empirical support for Levelt et al.'s model of speech production

Two issues put forth in favor of discrete serial models are resyllabification and incremental syllable construction. Resyllabification is the most intuitive of these arguments. We know that syllabification may occur across word boundaries. Levelt (1999) illustrates this point with the sentence *They will select* *us*. The syllabification of the final two words is /sI.lɛk.tʌs/. The final /t/ of the word *select* surfaces as the onset of the final syllable, rather than in the coda position of the second (as it would if *select* were pronounced carefully in isolation). If all metrical and syllabic structure is stored with the words in the mental lexicon, and there is no later level of phonetic encoding before articulation, as is the structure proposed by Dell's model, it is difficult to account for resyllabification processes. This is particularly prevalent in Spanish. While judgments of English syllabification may be blurred by issues such as ambisyllabicity, in Spanish syllable boundaries are very clear. There is obligatory resyllabification of coda consonants to the onset position of the following word if the second word begins with a vowel.

Meyer (1990, 1991) introduced the implicit priming paradigm to study the time course of phonological encoding. Her major finding was that a word's form is built up incrementally, starting with the first segment. In the implicit priming paradigm, participants first learn prompt-response pairs of words, e.g. *single-loner, place-local, fruit-lotus*. They are instructed to produce the response word when they see the prompt word. For example, if prompted with *single*, the participant should say *loner*. In homogenous sets, the response words always share a phonological property, such as the first syllable *lo-* in the examples above. In heterogeneous sets, the response words do not share the phonological property. The hypothesis is, if people are attuned to the implicit prime in the homogeneous groups, i.e. they can prepare the first syllable of the word in the sets in which the first syllable is always the same, then results should evidence faster reaction times, as measured in milliseconds from the presentation of the prompt word to the articulation of the response word. Of course, in the heterogeneous sets, there is no implicit prime. Therefore, the prediction would be that no priming effect should surface. Meyer (1990, 1991) shows that response latencies were significantly shorter in the homogeneous condition than in the heterogeneous condition.

In order to understand the time course of syllable encoding, Meyer conducted a second experiment in which the response words' first syllables did not match phonologically, but their second syllables did. In this experiment, no implicit priming effect was found. The presence of an effect when the first syllable matched, but not when only the second syllable matched, suggests that prosodification is constructed incrementally. That is to say, speakers cannot prepare for a later syllable until the first syllable is encoded. Roelofs & Meyer (1998) used a similar experimental paradigm in which homogenous/heterogeneous sets were matched/varied for stress placement to show that speakers cannot prepare for the first syllable unless they already know the word's metrical frame.

These findings offer strong support for Levelt et al.'s model in which metrical frames are stored with the word at the lemma level. In their model syllable encoding is assumed to occur online after the word has been selected from the mental lexicon. The data from these time course studies suggest that metrical information is accessed first, and syllabification/prosodification occurs at a later stage. The implicit priming paradigm appears to tap into a stage in processing where syllables emerge at the interface between phonological and phonetic encoding (Cholin et al., 2004). Additionally, syllables are encoded syllable-by-syllable in an incremental fashion. It is difficult to account for such findings within a connectionist framework in which all syllabic information is assumed to be stored with the lexical items and all syllables in a word are encoded in parallel. If this is the case, researchers should not be able to find effects of incremental syllable construction. In a network in which all syllable information is available at the same time, it should be possible to prime syllables from any position in the word.

1.4.2.5 Evidence of syllable encoding from Spanish

In Spanish, a number of studies have shown results that support syllabic processing during visual word recognition (Álvarez, Carreiras & de Vega, 2000; Álvarez, Carreiras & Taft, 2001; Álvarez, de Vega & Carreiras, 1998; Carreiras & Perea, 2002; Carreiras, Vergara & Barber, 2005). Two approaches include the standard priming paradigm and manipulations of syllable frequency. One main effect found is that words with high-frequency syllables produce longer response times in lexical decision and progressive demasking tasks than words with low-frequency syllables. This inhibitory effect of syllable frequency is interpreted in the literature as competition at the word level. Word identification is delayed for words with larger syllabic neighborhoods. A second approach to testing syllable encoding is the examination of effects within a priming paradigm.

Basing their study on an activation-based model in which sublexical input phonology is structured syllabically (cf. Ferrand, Segui & Grainger, 1996), Álvarez, Carreiras, & Perea (2004) ran a series of experiments employing a masked priming paradigm to test whether syllabic effects arise from a sublexical phonological level or from a sublexical orthographic level. In their first experiment, critical stimuli consisted of real words in Spanish that were preceded by nonword primes that always shared the first three letters. Half of the primes shared the first syllable with the target word (e.g. *ju.nas-JU.NIO*), and half did not (e.g. *jun.tu-JU.NIO*). The results showed a priming effect for prime-target pairs that shared the first syllable. These findings provide empirical support to the argument that, at least in Spanish, the syllable is an intrinsic part of word recognition processes. However, they do not tease apart the question of the locus of syllable encoding. That is to say, we can not tell from these data whether the effect appears due to facilitation at a sublexical phonological or orthographic level.

While the Spanish spelling system is highly regular and transparent, in which there tends to be a one-to-one correspondence between graphemes and the phonemes they represent, there are examples of ambiguity. For example, the graphemes "j" in all contexts and "g" followed by "i" or "e" both represent the phoneme /x/. Both "b" and "v" correspond to /b/. In experiment 2, Álvarez et al. (2004) took advantage of these grapheme-phoneme ambiguities in Spanish to investigate whether syllabic effects are phonological or orthographical in nature. As in the previous experiment, critical items were preceded by nonword primes.

However, in this case some primes shared the same orthographic and phonological initial syllable (e.g. *vi.rel-VI.RUS*) and some only shared phonology and differ orthographically (e.g. *bi.rel-VI.RUS*). Control primes were also included that shared the same number of letters with the targets but that did not share the initial syllable (*vir.ga-VI.RUS/bir.ga-VI.RUS*). The prediction here is that, if syllables are phonological units of processing, priming effects should be observed in the syllable-matched conditions, regardless of orthographic overlap.

Results show that CV words were responded to more quickly when they were preceded by a nonword with the same initial syllable than when they were preceded by a nonword with a different syllable. This effect was found both when the orthography matched as well as when it does not. This suggests that the syllabic priming effects are phonological in origin.

Álvarez et al.'s (2004) final experiment examines whether the phonological effect found in their second experiment is truly caused by the first phonological syllable, or if it may rather be due to the fact that the primer-target pairs shared a rime/body. While the results thus discussed underscore the importance of the full first syllable in Spanish, there is previous work in English that brings into question what part of the syllable is pushing the effect.

Chateau & Jared (2003) argue for BOB effects (body of the basic orthographic syllable structure effects), basing their work on Treiman, Mullennix, Bijelic-Babic & Richmond-Welty's (1995) work on orthographic body units and Taft's (1992) arguments for the importance of the BOSS (basic orthographic syllable structure). Treiman et al. (1995) found that the orthographic representation of the phonological rime constitutes an important level of consistency between the English orthographic system and phonology. At the grapheme-phoneme level, the English language appears very inconsistent. However, when the body units (referred to as orthographic rimes in Treiman et al. (1995)) are taken into consideration, English orthography becomes much more consistent (Lesch & Pollatsek, 1998). Treiman et al. (1995) conclude that adults are not limited to an alphabetic strategy in translating from spelling to sound, but rather they make judgments based on larger patterns of consistency between multiple-grapheme and multiple-phoneme units. As English orthographic vowels are much less consistent in their mapping to phonology than consonants, Treiman et al. (1995) provide statistical evidence that the consonants that follow the vowel are a much greater cue to the pronunciation of the vowel than those that precede it. Therefore, VC_2 units yield strong patterns of consistency to phonology. Cueing into these patterns is a potential way for readers to deal with "the vagaries of the English writing system" (Treiman et al., 1995, p. 130).

Adapting Treiman et al.'s findings, Chateau & Jared (2003) argue that it is not just the syllable structure that matters in disyllabic words. For example, the word *vertex* is divided syllabically into *ver.tex*. However, the orthographic sequence *–ert* is a possible word ending. The fact that the grapheme "e" is followed by two consonants, regardless of the syllabification of those segments, is a strong cue to the phonological mapping of the grapheme "e". It has been found that consistency effects exist and are more reliable based on these possible word endings, even within words, rather than purely based on syllable structure. Chateau & Jared (2003) argue that, given that 83% of English words with two syllables are stressed on the first syllable, readers can assume that the second vowel in a disyllabic word is lax. Therefore, it is advantageous to pay close attention to the consonants that follow the first vowel, as they are strong cues to how the first vowel should be pronounced. The attention paid to these segments leads to patterns of activity in the network and cause consistency effects. Therefore, at least in English, there is evidence that speakers cue into information in the orthography beyond the basic level of the syllable.

To test whether this is the case in Spanish, in their last experiment Álvarez et al. (2004) devised a "rime-only condition" that was tested against the phonological-syllable priming condition previously described. For example, *va.lis-BA.LÓN* is a prime-target pair that is matched phonologically across the entire first syllable (the graphemes "v" and "b" represent the same phoneme in Spanish, /b/). Pairs such as this one were compared to pairs that were matched only for the rime of the first syllable (e.g. *fa.lis-BA.LÓN*). The final experiment also contained a condition in which the first three phonemes matched, but their syllabification differed (e.g. *val.ti-BA.LÓN*).

The results for Experiment 3 find that response times are substantially faster when primes and targets shared the first phonological syllable than when primes and targets only shared the first three phonemes or when they only shared the rime/body of the initial syllable (Álvarez et al., 2004). They also found a clear superiority of the syllable over the rime which supports the syllable as a sublexical unit in Spanish. Their findings offer convincing empirical evidence that syllables are fundamental units in the processing of Spanish words and the syllabic effects observed are phonological (as opposed to orthographic) in nature. They also suggest that there are different reading strategies across languages. The Spanish data indicate that native speakers of Spanish pay close attention to the full syllable when segmenting words, whereas previous research shows that English speakers take advantage of orthographic information beyond the first syllable. Given the absolute correspondence between Spanish orthography and phonology for vowels, native speakers of Spanish do not need to look for cues in the orthography beyond the vowels in order to decide which phoneme the vowel grapheme represents. Therefore, at least in Romance languages, the syllable seems to act as the most relevant sublexical unit of processing (Álvarez et al., 2004; see Ferrand, Segui, & Grainger, 1996, for evidence of syllabic priming in French).

Additional evidence in support of a syllable level of encoding within psycholinguistic models comes from a body of literature discussing effects of syllable frequency. In their 1994 study, Levelt and Wheeldon tested the syllabary hypothesis by examining the theory's prediction of a syllable frequency effect. Accessing a syllable in the mental syllabary that is frequently used in the language will be faster than accessing a syllable that is less frequently used. Therefore, the authors devised a set of stimuli that consisted of four types of words: low-frequency words with low-frequency syllables, low-frequency words with high-frequency syllables, high-frequency words with low-frequency syllables, and high-frequency words with high-frequency syllables. Subjects first underwent a study phase in which they learned to associate symbols with disyllabic response words. In the second phase of the experiments, the subjects were presented with the symbols and produced the corresponding word. Results showed both word and syllable frequency effects. Higher word and syllable frequencies produced faster response times. Crucially, however, the two effects proved independent of each other. These findings suggest that the mental lexicon and the mental syllabary are indeed distinct levels of representation in the mind.[22]

In a later study, Ferrand, Segui & Grainger (1996) argued for an interactive activation model of visual word recognition and naming with the presence of syllable-sized units both in the sublexical input phonology and in the sublexical output phonology. They hypothesized that the syllable-sized units in the output phonology could facilitate articulatory responses. In order to test their hypothesis, the authors employed a masked priming paradigm across four tasks: word naming, nonword naming, lexical decision, and picture naming. Interestingly, they found facilitative effects of syllable priming for real word, nonword, and picture naming, but no syllable effect was obtained in the lexical decision task. Therefore, the authors argued that the syllable priming effects observed are located at the level of output phonology. That is to say, it is within the processes of articulation that syllable priming effects occur.

---

[22] For corroborative evidence obtained with ERP data, see Carreiras, Vergara & Barber (2005).

Framing their series of experiments on these two studies, Carreiras & Perea (2004) examined issues of syllable frequency in Spanish. They observed that the previous literature testing syllable effects with masked priming tasks manipulated the first syllable of the prime to match, or not, the first syllable of the target stimulus. However, Levelt & Wheeldon (1994) obtained a syllable effect only for the second syllable of their disyllabic items. Therefore, in their series of experiments, Carreiras & Perea (2004) set out to examine the issue of syllable frequency effects and which syllable evidences these effects in Spanish.

In their first experiment, Carreiras & Perea (2004) manipulated the frequency of both the first and second syllables of 84 disyllabic nonwords. In a standard naming task, they found a facilitative effect of syllable frequency only for the initial syllable. They also found no signs of any effect of syllable frequency for the second syllable. One possible confound to this finding is that disyllabic nonwords with a CV.CV syllable structure (as their stimuli were) are always read with penultimate stress in Spanish. Therefore, it is possible that the effect obtained was the result of stress placement rather than purely syllable frequency. To test this hypothesis, the authors conducted a second experiment in which all stimuli were diacritically marked for final stress. Again, results find a syllable frequency effect for the first syllable but not for the second syllable. There was also no sign of an interaction between lexical stress and syllable frequency. Carreiras and Perea argue that these findings are consistent with Levelt's model, and that the process of stress assignment and the process of retrieving syllables from the syllabary seem to occur at different stages of processing (Carreiras &

Perea, 2004). The data to be presented in this dissertation challenge the independent nature of stress placement and syllable structure and suggest a much closer interaction between the two than argued by Carreiras and colleagues in Spanish or Levelt in English and Dutch. The data presented in Carreiras & Perea (2004) also differ drastically from Levelt & Wheeldon's (1994) findings in that in Spanish it appears to be the first syllable that causes syllable frequency effects.[23] Nevertheless, these data offer strong support for the concept of the mental syllabary and the need for a separate locus of syllable encoding within psycholinguistic models. This dissertation examines further the nature of syllable-level encoding and its relationship with stress placement via behavioral data from Spanish.

1.5 Structure of the dissertation

The central approach taken in this dissertation is inherently cross-disciplinary. Building on the linguistic and psycholinguistic work reviewed in the previous sections of this chapter, this dissertation explores issues of phonotactic knowledge in Spanish from a novel perspective by adapting psycholinguistic methodologies as a means of informing both models of phonology and models of speech production. More specifically, this research examines the interaction of stress assignment and syllable structure in Spanish. If native speakers of Spanish

[23] The syllable frequency effect was also found to be independent of bigram frequency in a third experiment not discussed here.

cue into the differential behavior of rising and falling diphthongs, then a stochastic perspective of the lexicon would predict that, when tested in a behavioral paradigm, speakers will react differently to proparoxytones deemed theoretically impossible due to the presence of a rising or falling diphthong. Such results would challenge traditional theories of syllable weight which call for a strict binary distinction between light and heavy syllables. If we assume Spanish is a weight sensitive language, then traditional accounts would argue that both rising and falling diphthongs form heavy syllables. Empirical findings of a distinction between the two would be evidence in favor of gradience in syllable weight. This dissertation collects behavioral data from native speakers to investigate these questions in order to inform Spanish phonotactics and how phonotactic knowledge is represented in phonological theory in general.

The dissertation also informs psycholinguistic models of speech processing. To that end, this thesis consists of a series of experiments utilizing variations on the naming task to probe for the locus of stress and syllable structure encoding. Specifically, comparisons of results from a series of speeded naming tasks and a delayed naming task speak to models of speech production as to the level of processing in which stress and syllable structure interact.

1.5.1 Description of behavioral tasks

The combination of tasks was chosen because they lend themselves well to probing various levels of processing under a behavioral paradigm. In contrast with intuition-based linguistic methodology, these tasks allow the researcher to collect native speaker reactions to controlled stimuli with little to no metalinguistic introspection. The experiments consist of asking participants simply to read words on a screen as quickly and accurately as they can, yielding data in the form of reaction times and error rates across a range of critical stimulus categories.

The naming task is a particularly appropriate task for a study of this nature in that it forces phonological activation. Given that the focus of this study is phonotactic knowledge, requiring actual production is the most natural means by which to test the processing of such knowledge. The naming task is well established and frequently employed in psycholinguistic research (Andrews, 1997; Andrews & Heathcote, 2001; Levelt, 1999; Stemberger, 2004; Dell, Lawler, Harris & Gordon, 2004; among many others). The standard speeded naming task consists of the presentation of a fixation point in the middle of a computer screen, followed by a short delay, and then the presentation of the stimulus. When the word appears, participations read what they see aloud. Reaction time data is collected in milliseconds. Recordings may also be made for subsequent coding for accuracy. This task is employed in the first three experiments of this dissertation.

The task utilized in the final experiment in the series is a delayed naming task. This methodology arguably serves to remove any lexical effects by giving adequate time for lexical processing. As in the standard naming task, stimuli are presented after a fixation point on the screen. However, instead of instructing participants to produce what they see immediately, they are instructed to withhold articulation until prompted to speak (e.g. when they hear a tone, when the word changes color, etc.). The delay between the presentation of the stimulus and the prompt varies. Longer delays allow time for any lexical processing to occur. Therefore, any remaining effects to be found are attributed to the production/articulation level of processing (Andrews & Heathcote, 2001).

In sum, the comparison of these two naming paradigms forms a research paradigm that is able to inform models of speech production as to the level of processing in which stress placement and syllable structure constraints interact.

1.5.2 Subsequent chapters

The remaining chapters of this dissertation proceed as follows. Chapter 2 tests the methodology by examining various phonotactic constraints on the Spanish three-syllable window within a behavioral paradigm. It also discusses the issue of historical gaps and synchronic prohibitions on stress placement and syllable structure in the phonological grammar of modern Spanish. Chapter 3 builds on the work discussed in Chapter 2 by focusing particularly on rising and falling diphthongs to test for differential patterning between the two diphthong types. This chapter also discusses the linguistic import of the results as concerns theories of syllable weight and gradience in the phonological system. Chapter 4 replicates the results presented in Chapter 3 with a different population (i.e. a different dialect group within the Spanish speaking world). In Chapter 5, I present the results of the delayed naming task and discuss implications for the locus of phonological encoding in models of speech production. Finally, I close the thesis in Chapter 6 with a general discussion of the findings.

Chapter 2

Experiment 1: Speeded Naming Task - General

2.1 Introduction

Experiment 1 seeks tests native speaker knowledge of phonotactic patterns via the collection of behavioral data. Specifically, this experiment examines how native speakers process the three-syllable window for Spanish stress assignment, by probing the gradient nature of phonotactic knowledge from a psycholinguistic perspective.

As the issue under question is of a phonological nature, the naming task provides a useful methodology in that it forces phonological activation. This task is a novel approach to the study of the Spanish stress system in that it does not depend on native speaker intuitions which rely heavily on metalinguistic interpretation and introspection. Rather, the naming task provides a more direct examination of the cognitive processes native speakers employ when they prepare words for speech. As opposed to pure description or survey-style data collection, the naming task collects data in the form of accuracy and latency in the common context of reading words aloud.

In order to test the boundaries of the phonological system, i.e. what *cannot* exist in Spanish as well as what can and does exist, Experiment 1 tests native speaker acceptability of various phonotactic restrictions to the three-syllable window by presenting participants with nonword stimuli. Critical experimental items consist of word forms based on theoretically prohibited or unattested phonotactic patterns. The specific violations of phonotactic constraints are discussed in detail under the Materials section of this chapter. The prediction is that participants will commit more production errors and/or take longer to produce forms that are not permitted by the synchronic phonological grammar. Therefore, the collection of accuracy and latency data from proscriptions (forms deemed prohibited by the grammar) and gaps (forms that are unattested although phonologically licit) offers empirical evidence to support or refute linguistic accounts of syllable weight and of the interaction between Spanish stress assignment and syllable structure. This methodology allows us to examine the cognitive representation of phonotactic knowledge under a nuanced, finely-grained lens.

2.2 Participants

32 functionally monolingual adult speakers of Spanish participated in Experiment 1. All were undergraduate students at the University of Jaen, Spain. One participant's data were excluded due to problems with the digital recording.

2.3 Materials

The critical stimuli for this experiment are based on four attested restrictions on the Spanish three-syllable window (cf. Roca, 1991). Three-syllable nonwords were designed with the following characteristics:

Table **1**: Categories for experimental items in Experiment 1

1. penult closed by a consonant (e.g. *dóvalda*)
2. "n" as the final consonant (e.g. *dólagan*)
3. rising or falling diphthong in the penult (e.g. *dóbiana*)
4. final syllable with a palatal onset (e.g. *dóvaña*)

The first category, in which the penultimate syllable contains a consonant in coda position, exhibits a syllable structure which only permits final or penultimate stress assignment (Harris, 1983, 1992; Lipski, 1997). Proparoxytones with a closed penult, with rare exceptions such as foreign borrowings like *Wáshington*, are unattested in the Spanish lexicon. Therefore, if a nonword such as *jomaldo* is presented in a naming task, the participant should have little difficulty pronouncing the word. However, if the stimulus is marked diacritically for antepenultimate stress, such as in *jómaldo*, phonological theory predicts that this prohibited combination of stress and syllable structure should cause problems for the participant. In a naming task, the prediction is that speakers will either take longer to pronounce such a word, commit more errors when pronouncing them, or a combination of both.

Closed final syllables also have been observed to shrink the three-syllable window to one of two syllables. There are, however, many words in Spanish that end in consonants with antepenultimate stress, e.g. *régimen, espécimen, Júpiter, máximum, déficit*, etc. (Roca, 2005). Many of these words are *cultismos,* that is, words that have been borrowed into Spanish directly from Latin without the effects of natural phonological evolution. As discussed above, Spanish treats verbs and "nonverbals" differently when assigning stress (Roca, 2005, 1991; Harris 1992, 1983). While the large majority of consonant-final nonverbals carry final stress (97.8% according to Núñez Cedeño & Morales-Front, 1999), almost all third-person plural verbal forms end in *–n* with penultimate stress.[24] Presumably due to these forms, and the presence of final *–s* either to mark plurality or the second-person singular verbal form, Spanish orthography groups words ending in orthographic *–n* or *–s* in the same category as words that end in a vowel, i.e. words whose unmarked pattern is penultimate stress. Therefore, a dichotomy exists between verbals and nonverbals ending in *–n*. Verbal forms prefer penultimate stress, whereas nonverbals evidence a strong preference for final stress. Spanish orthography marks a preference for the former. In both cases, antepenultimate stress seems to be highly restricted. However, the exceptions cannot be ignored, as forms such as *régimen*, are far from infrequent in the language.[25] The prediction is that the general ban on proparoxytones

---

[24] The future tense, which carries final stress, is the exception. The Modern Spanish future tense is the synthetic result of the periphrastic future tense in Latin, i.e. infinitive + HABEO.
[25] The Spanish *régimen* may be translated into English as "diet" or also as "regime", frequently used to refer to Spain under Franco's rule.

ending in consonants will suffice to evoke substantial production errors and/or longer reaction times when participants are forced to read aloud forms such as *cónabon* or *pómagan*. The stimuli designed for this category all end in –*n*. If the prediction holds, even for words that evidence the pattern that most often presents exceptions (i.e. –*n* final proparoxytones), then this would be strong evidence in favor of the psychological reality of the descriptive rule expounded in traditional accounts.

The issue of rising and falling diphthongs and their interaction with stress placement has been discussed previously in section 1.3.4 of chapter 1. It is clear from the literature that there is a strict ban on proparoxytones with a diphthong in the penultimate syllable. Words such as *lótiago* or *góvaino* are simply unattested in the Spanish lexicon. Therefore, in a naming paradigm word shapes such as these are predicted to cause higher latency and lower accuracy in participant responses.

The last category consists of three-syllable words with a palatal onset in the ultima. The lack of words in Spanish with this structure can be understood as a historical residue from Latin. With the exception of borrowed words, palatals in Modern Spanish all derive from phonological change, as the Latin phonemic inventory contained no palatal consonants. Palatals in the onset of the final syllable typically derived from the merger of two Latin phonemes (Penny, 2002; Lloyd, 1987). Modern Spanish palatal consonants can be traced to four principle historical changes: gliding of atonic Latin E and I in hiatus, syllable-initial velars, syllable-final velars, and the lenition of Latin LL and NN.

Atonic E and I in hiatus with other vowels in Latin frequently underwent a process of lenition in which they were reduced to the palatal glide [j]. This change produced two subsequent changes. First, the gliding of the vowel results in the loss of a syllable. Second, in many instances the newly-formed palatal glide palatalized the preceding consonant, as illustrated by the following examples:

Table **2**: Deriving palatal final onsets from atonic Latin E and I in hiatus

ARĀNEA (a.ra.ne.a) > a.ra.nja > a.ra.ɲa (*araña*)

HISPANIA (his.pa.ni.a) > es.pa.nja > es.pa.ɲa (*España*)

PODIU (po.di.u) > pɔ.ðjo > po.jo (*poyo*)

The result of the gliding of the Latin atonic E and I results in a palatal onset for the final syllable.

Syllable-initial voiced velars also evolved to palatal consonants in Modern Spanish. Initial G in Latin is hypothesized to have had a fronted pronunciation before front vowels which became the voiced palatal fricative [ʝ]. This change brought about a merger in pronunciation with word-initial atonic I. This pronunciation was maintained in words such as the following (written *y* in Old Spanish):

Table **3**: Deriving palatal initial onsets from syllable-initial velars in Latin

GEMMA (gem.ma) > gɛm.ma > je.ma (*yema*)

GYPSU (gip.su) > je.so (*yeso*)

Velars in syllable-final position also underwent a process of lenition in which they first became a velar fricative and later a palatal glide. These coda-position glides frequently palatalized the onset of the following syllable, combining with it.

Table **4**: Deriving palatal final onsets from syllable-final velars in Latin

FACTU (fak.tu) > faj.to > fej.to > hej.to > e.t͡ʃo (*hecho*)

LIGNA (lig.na) > lej.na > le.ɲa (*leña*)

A similar pattern appears with the sequence ULT in Latin. Syllable-final L was often velar in Vulgar Latin. This velar quality was exaggerated by the high back vowel /u/ and, therefore, underwent a process similar to the previous syllable-final velars.

Table **5**: Deriving palatal final onsets from Latin - ULT

MULTU (mul.tu) > moj.to > mu.t͡ʃo (*mucho*)

The last category of phonological change to produce the Modern Spanish palatals are the Latin geminates LL and NN. The Latin consonantal inventory underwent a systematic change in which voiceless geminate stops became singleton voiceless stops, voiceless stops became voiced stops, and voiced stops became voiced fricatives as illustrated below:

Table **6**: Lenition of Latin geminates

/pp/ > /p/	/tt/ > /t/	/kk/ > /k/
/p/ > /b/	/t/ > /d/	/k/ > /g/
/b/ > /β/	/d/ > /ð/	/g/ > /ɣ/

However, if the geminates LL and NN had followed the same pattern of lenition, they would have become the singletons /l/ and /n/. Unlike the voiceless stop singletons in Latin that underwent voicing, the extant L and N were already voiced and typically suffered no other form of lenition. This process would have resulted in the neutralization of LL and NN with L and N. Apparently to avoid this situation, the geminates LL and NN underwent a process of palatalization instead.[26]

Table **7**: Deriving palatal final onsets from the Latin geminates LL and NN

CABALLU (ka.bal.lo) > ka.βal.lo > ka.ba.ʎo (*caballo*)

ANNU (an.nu) > a.ɲo (*año*)

For the purpose of this study, what is important is that the formation of Spanish palatal final onsets resulted from the loss of a segment from the original

[26] In most dialects of modern Spanish the palatal liquid has been replaced with a palatal glide, fricative, or affricate.

Latin. The palatal segment derives either from the loss of the Latin penultimate vowel, or from the merger of the penultimate coda consonant and the original final onset. Therefore, the antepenultimate syllable of any modern Spanish word with a palatal in the onset position of the ultima would have been either the preantepenultimate syllable (in the case of palatals created through vowel loss), or the antepenultimate syllable of a word with a heavy penult (the case for palatals created through consonant loss).

The phonological system of Modern Spanish theoretically should hold no synchronic prohibitions on proparoxytones with the syllable structure CV.CV.PV (where P represents a palatal consonant). In the current state of the language, all of these syllables are light and antepenultimate stress should be possible. However, due to the diachronic reasons previously discussed, these forms do not exist. They thus represent a gap in the Spanish lexicon.

To summarize, the following table gives example stimuli for each of the critical categories with quantities in parenthesis.

Table **8**: Stimuli summary for Experiment 1

Critical Items	*Antepenultimate*	*Penultimate*
Closed Penults	dóvalda (14)	dovalda (14)
"n" as final consonant	dólagan (14)	dolagan (14)
Rising diphthong	dóbiana (14)	dobiana (14)
Falling diphthong	dótaiga (14)	dotaiga (14)
Palatals - Singleton	dóvaña (14)	dovaña (14)
Palatals - Digraph	dóvacha (14)	dovacha (14)
CV.CV.CV. Controls	dóvasa (28)	dovasa (28)

All stimuli were presented both with and without a diacritic accent mark over the first syllable. According to Spanish orthographic rules, for any word to be stressed as a proparoxytone it must be marked diacritically. Any word that ends with a vowel will be assigned penultimate stress unless marked diacritically to be stressed otherwise. However, no two participants saw the same stimulus in both the antepenultimate and penultimate condition. Instead stimuli were presented in a counterbalanced fashion in which the first participant saw *dóvalda*, the second participant *dovalda*, and so forth. This was done to avoid any possible effects the repetition of the segmental sequence might cause.

The segmental composition of the stimuli was also strictly controlled. When developing the stimuli, special attention was paid to avoid any possible intraword lexical items. For example, the nonword *gotapa* could not be chosen as a control item because it contains the existing Spanish words *gota* and *tapa*.

The total 112 experimental items were presented randomly with 140 real words. These filler stimuli were chosen from the Cuetos & Alameda (1995) database to form two groups of high and low frequency items further divided into groups of antepenultimate, penultimate and final stress. A comprehensive list of the stimuli for this experiment is found in Appendix A.

2.4 Procedure

Participants were tested individually in a quiet room. They were seated in front of a computer screen and two microphones. The first microphone was connected to a PST Serial Response Box, and the second to a digital recorder (Marantz PMD660 sampling at 48 kHz). The button box served as a voice key to cue the computer to record reaction times and to present the next stimulus when the participant spoke. The digital recorder was used to record the experimental session for subsequent coding of the participants' responses for accuracy.

In this experiment the participants performed a standard naming task. First, an instructions screen informed the participants that they would be reading words one by one on a computer screen, and that they should read the words aloud as quickly and accurately as possible. Following the instructions screen, participants performed a practice session with 12 items to become accustomed to the task.

Each trial began with a fixation point (+) which was presented in the middle of the screen in black font on a white background. This was done to focus the participants' attention to the center of the screen and to cue them that the word they needed to pronounce was about to appear. The fixation point remained on the screen for 500 milliseconds and was followed by a blank screen for 500 milliseconds. Following these 1000 milliseconds, the stimulus was presented and remained on the screen until the participant spoke. The participants' speech cued the voice key, which caused the stimulus on the screen to disappear and to be replaced with a blank white screen. In the case that the participant committed an error, during the blank screen the experimenter cued the computer to present an error screen. The word "Error" appeared in red letters in the upper left-hand corner of the screen. The purpose of this error message was to alert the participant to the fact that they had committed an error and to focus them on the task. After the error message, another blank screen appeared for 1000 milliseconds until the following trial began.

Once the participants completed the practice trials, the experimental stimuli were presented in random order over two experimental blocks with a 30 second break in between them. The following diagram exemplifies the naming task as it was employed for this experiment.

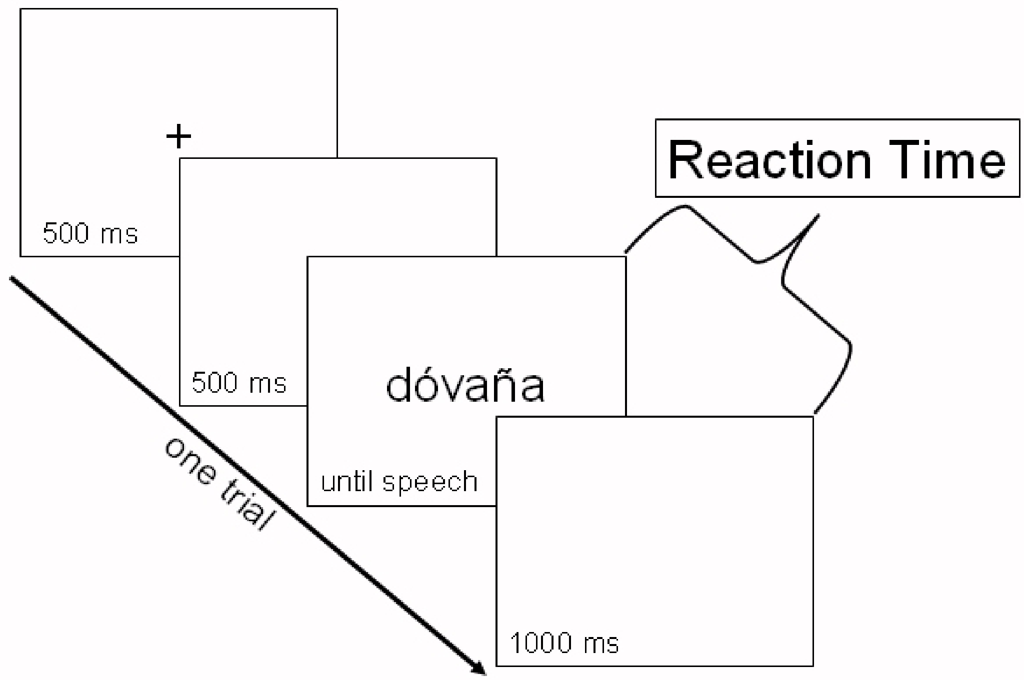

Figure **10**: Procedure for Experiment 1 (standard naming task)

2.5 Data Analyses

Both error rate analyses and reaction time analyses were conducted on the data from this experiment. Native or near-native listeners (including the author) coded the errors manually using digital recordings of the experimental sessions. Errors comprised two categories: production errors and technical errors. Production errors consisted of instances in which participants misspoke the stimulus. Examples of this type of error were stress shifts to another syllable, changes in the segmental makeup of the word, or disfluencies in which words were not finished or contained pauses. Technical errors were those in which the microphone or the response box failed. These types of errors resulted in either extremely long reaction times, such as when the microphone did not trigger the voice key after articulation, or extremely fast reaction times, such as when the voice key triggered the next stimulus prematurely before the participant actually spoke.

The data were cleaned as follows. First, technical errors due to the microphone and/or voice key were eliminated from all analyses. Production errors were included in error rate analyses, and excluded from reaction time analyses. All reaction time latencies were cleaned to exclude any responses faster than 200 ms and slower than 2000 ms. Subsequently, any latencies 2.5 standard deviations above or below each individual's mean reaction time were also eliminated.

Both reaction time and error data were submitted to separate repeated measures ANOVAs by participants (*F1*) to test for variance across the population and by items (*F2*) to test for variance across the stimuli set. Main effects and interactions were followed up by Tukey's HSD post-hoc tests to identify any independent effects across the experimental categories and conditions. Results are discussed below.

2.6 Results and Discussion

2.6.1 General findings

Analyses of error data reveal a significant main effect of word type [$F1(4, 120) = 73.38$, $MSE = 1.76$, $p<.001$; $F2(4, 107) = 50.22$, $MSE = 1.07$, $p<.001$], a main effect of stress type approaching significance by participants and significant by items [$F1(1, 30) = 3.62$, $MSE = 0.33$, $p<.067$; $F2(1, 107) = 11.62$, $MSE = 0.22$, $p<.001$], and a significant interaction of word type and stress type [$F1(4, 120) = 119.78$, $MSE = 2.26$, $p<.001$; $F2(4, 107) = 66.42$, $MSE = 1.24$, $p<.001$]. As the focus of this experiment is to examine the effect of syllable structure phonotactics on stress placement, the important finding is the significant interaction of word type and stress type. These results suggest that there is a different effect of syllable structure in different stress conditions. That is to say, while it is true that participants react differently to varying word types, more importantly, their treatment of each word type is modulated by the stress assigned to the word.

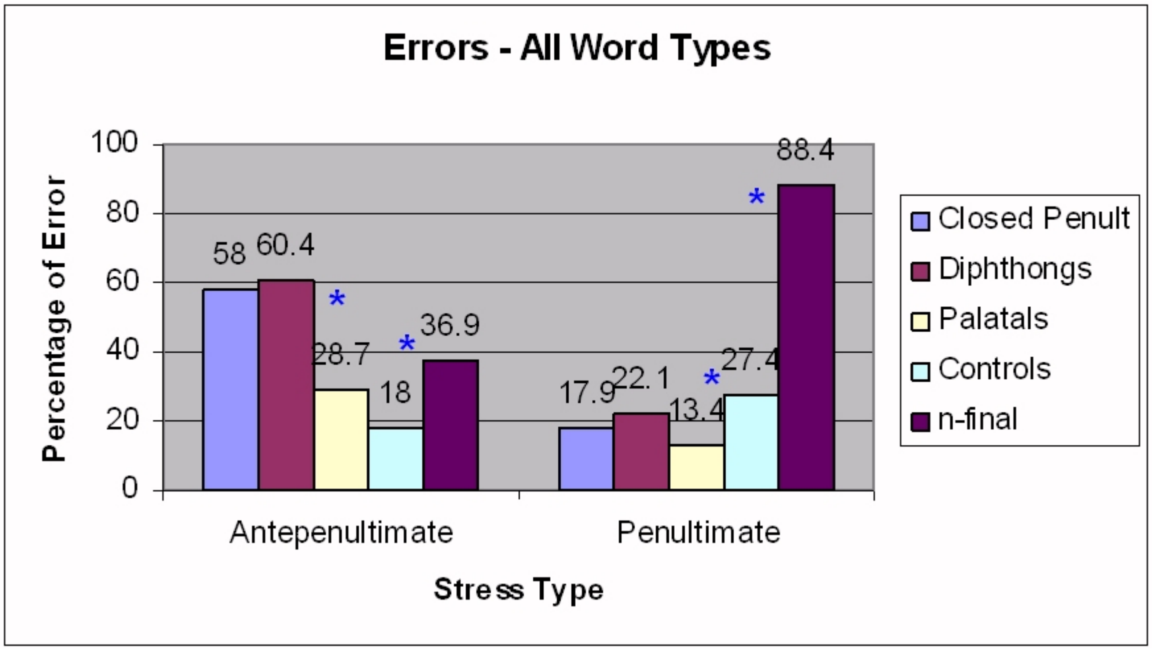

Figure **11**: Experiment 1: Error Rates – All word types

Figure 11 represents the percentage of errors participants committed on each word type.[27,28] The antepenultimate condition represents stimuli that were presented with a diacritic over the first syllable to designate proparoxytone stress. Stimuli presented without the diacritic (denoting paroxytone stress) are represented in the penultimate condition.

Perhaps at first glance the most striking result illustrated in the graph is the overwhelming amount of errors in the n-final category of nonwords in the penultimate condition. This result implies that native speakers are sensitive to

---

[27] Note that for the purposes of this analysis, the singleton and digraph palatal categories were collapsed into one palatal category, based on analyses that revealed no effect of grapheme length [$F1(1, 30) = 0.27$, $MSE = 0.00$, $p>.60$; $F2(1, 26) = 0.11$, $MSE = 0.00$, $p>.75$].

[28] For this initial analysis, the rising and falling diphthongs were also combined into one category. This was done because theories of syllable weight treat both diphthong types equally. Our prediction of differential behavior between the two diphthong types is discussed in a separate analysis below.

the fact that almost all nonverbals ending in consonants in Spanish are oxytones (cf. Núñez Cedeño & Morales-Front, 1999), irrespective of the prescriptive rules that require a paroxytone stress pattern if the final consonant is an *n*. For example, when presented with stimuli such as *pomagan* or *tanagon*, participants ignored the orthographic convention of penultimate stress in favor of a final stress pattern. This tendency was so strong that it provoked participants to err 88.4% of the time. In the antepenultimate condition, participants committed fewer errors in the n-final condition. This is arguably due to the presence of the diacritic. This marking draws the participants' attention to the first syllable, leading them to antepenultimate stress. When no diacritic was present, nothing focused the participants on a particular stress pattern. In this condition, the robust statistical preference for final stress surfaced.

Upon examination of the antepenultimate condition, it is clear that participants committed the highest amount of errors when presented with stimuli that contained either a closed penult or a diphthong in the penultimate syllable. These two categories pattern together, as post-hoc tests reveal no significant difference between them ($p>.99$). When participants were presented with theoretically prohibited phonotactic patterns, they committed production errors 58-60.4% of the time. The high percentage of error is indicative of the proscribed nature of proparoxytones with either a coda consonant or diphthong in the penultimate syllable. This conclusion is supported by the fact that participants committed substantially fewer errors in the control condition (58/60.4% vs. 17.9/22.1%). Assuming that the CV.CV.CV control items are licit forms in both the antepenultimate or penultimate conditions, and the forms are nonwords, it is expected that they will provoke a certain level of error. However, this level should be significantly lower than the predicted proscriptions, and this is exactly the finding obtained in Experiment 1 ($p<.001$).

In order to test the validity of the hypothesis that these results are due to violations of phonotactic constraints in interaction with stress placement, and not purely an effect of word type, we must also examine the behavior of these categories in the penultimate condition. If the phonotactic categories pattern similarly in the penultimate condition, one might argue that certain phonotactic constructions are simply more difficult to pronounce, regardless of the stress pattern assigned to the word. However, if these categories are treated equally in the theoretically licit condition (with penultimate stress), and differently in the theoretically prohibited condition (with antepenultimate stress), then this finding would offer strong support for traditional accounts of the three-syllable window in Spanish and would be consistent with the argument in favor of quantity sensitivity in Spanish. Post-hoc tests from Experiment 1 show no significant difference between the closed penult, diphthong, and control conditions in the penultimate condition.

Reaction time data were also submitted to 2 (stress type) x 5 (word type) repeated measures ANOVAs by participants and items. Results reveal no significant main effect of stress type [$F1(1, 11) = 1.67$, $MSE = 18735$, $p>.22$; $F2(1, 103) = 1.23$, $MSE = 6718$, $p>.27$], a main effect of word type that is significant by participants and approaches significance by items [$F1(4, 44) = 8.51$, $MSE =$

30050, $p<.001$; $F2(4, 103) = 2.40$, $MSE = 14505$, $p>.05$], and no significant interaction between stress type and word type [$F1(4, 44) = 2.20$, $MSE = 14941$, $p>.08$; $F2(4, 103) = 1.46$, $MSE = 7963$, $p>.21$]. These results indicate that participants treated the five word categories differently. However, the difference in stress placement did not appear to affect participant behavior. The following figure depicts participants' reaction times to the various word types in both stress conditions.

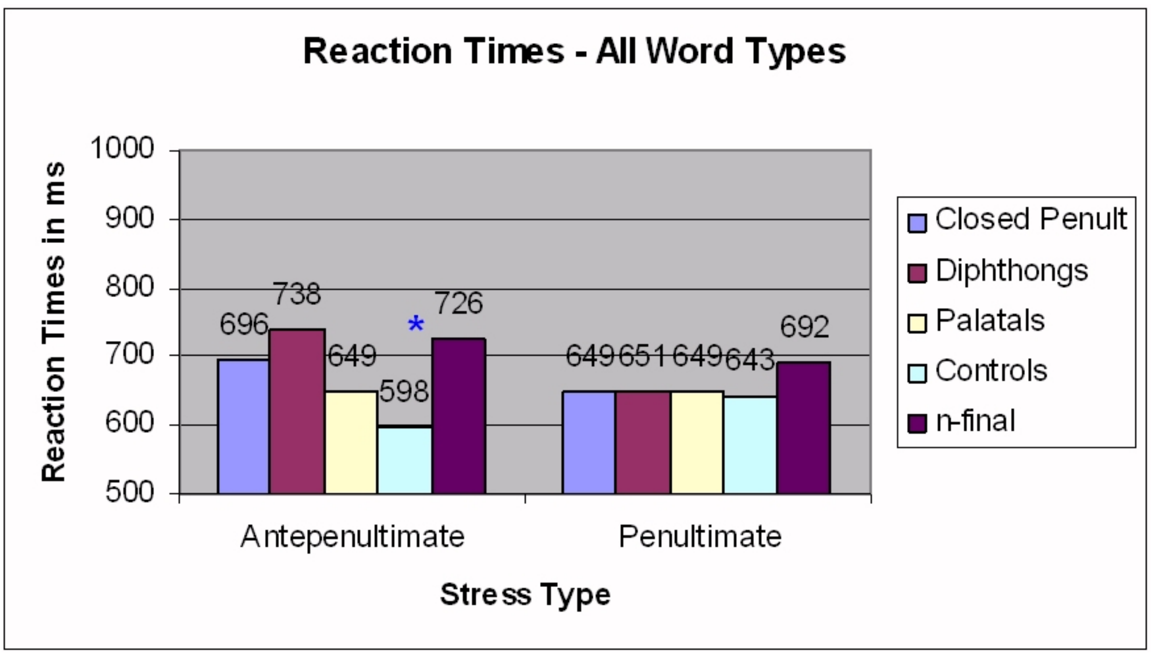

Figure **12**: Experiment 1: Reaction Times – All word types

The only differences that prove significant in post-hoc tests are between the diphthongs and the controls, and between the n-final and the controls, in the antepenultimate condition ($p<.02$). The differential treatment of proscribed vs. licit forms obtained in the error rate data are not confirmed in the reaction time data. The lack of an effect among the latency results is easily explained. The presentation of highly prohibited phonotactic sequences in a nonwords naming task produces an extremely high level of error rates. This is apparent in the accuracy analysis discussed above. The exclusion of inaccurate trials from the reaction time data renders empty cells. That is to say, there were participants who committed errors every time a stimulus of a certain category was presented. This poses a large problem for the analysis of the correct trials. Of the 31 total participants in this experiment, 19 had to be excluded from the latency analysis. The analysis of only 39% of the original data is too restrictive for the predicted results to be obtained. Given the nature of the task, it is evident that the error rate analysis is the more appropriate tool through which we can examine the effects of the interaction of phonotactic constraints and stress placement in Spanish.

In summary thus far, the findings of this experiment suggest differential representation of phonotactic constraints in the form of a division between proscriptions and gaps. Closed penults and glides in the penultimate syllable provoke significantly more errors in the antepenultimate condition than phonologically licit, although unattested, controls. Thus, it appears from these data that native speakers in Spanish do indeed differentiate cognitively between two types of forms that are absent from the modern Spanish lexicon: gaps, those words that could exist phonologically but do not; and proscriptions, words that are prohibited due to violations of phonotactic constraints in the synchronic grammar.

2.6.2 Palatal final onsets

Thus far in the discussion, I have ignored the behavior of the palatal final onsets. Perhaps the most intriguing finding in these data, however, lies in the behavior of the palatal stimuli. These items represent a pattern that is absent from the Spanish lexicon for purely historical reasons (see section 2.3 above). At the syllable level of generalization, the palatal stimuli are identical to the control stimuli. That is to say, both the controls and the palatal items exhibit a CV.CV.CV syllable structure. Thus, we expect the synchronic phonology of modern Spanish to treat both categories equally with regards to syllable weight. Both the stimuli with palatal final onsets and the control items should pattern as *light*. Light syllables are theoretically permitted in either stress condition. However, the results depicted in figure 11, suggest differential behavior between the two categories. In the antepenultimate condition, the palatal stimuli appear to provoke errors at an intermediate level between the true synchronic proscriptions (closed penults and diphthongs in the penultimate syllable) and the fully licit control items. The difference in error rates between the palatal items and the diphthongs and closed penults proves statistically significant in post-hoc pairwise comparisons ($p<.001$). As predicted by theories of syllable weight, the difference between the palatals and the controls is not significant ($p<.07$). However, descriptively, it appears that the palatal final onsets do provoke more errors. Also, in the penultimate condition, the palatal stimuli do indeed pattern with the proscribed forms in that there is no statistical difference between the palatal final onsets and the closed penults ($p>.95$), or between the palatal final onsets and the stimuli containing diphthongs (*p*>.27) when they are presented as paroxytones. A significant difference between the palatal stimuli and the control items surfaces in the penultimate condition as well (*p*<.01). The higher amount of errors for the control items is explained by the fact that they are completely licit forms in either the antepenultimate or penultimate conditions. The lower rate of error in the penultimate condition for the palatal stimuli implies the palatal items are less likely to allow mispronunciation favoring proparoxytone stress.

Building upon these observations, the data from Experiment 1 were submitted to additional 2 (stress type) x 3 (word type) repeated measures ANOVAs by participants and items to compare only the behavior of the closed penults, palatal final onsets and control items. This more specific analysis compares the palatal final onsets, the category of interest, with the two categories that represent the *heaviest* (closed penults) and *lightest* (controls) conditions, as predicted by phonological theory and suggested empirically in figure 11.

The second ANOVA revealed a significant main effect of word type [$F1(2, 60) = 23.59$, $MSE = 0.54$, $p<.001$; $F2(2, 67) = 15.77$, $MSE = 0.29$, $p<.001$], a significant main effect of stress type [$F1(1,30) = 21.65$, $MSE = 1.10$, $p<.001$; $F2(1, 67) = 36.76$, $MSE = 0.74$, $p<.001$], and most importantly a significant interaction of word type by stress type [$F1(2,60) = 68.04$, $MSE = 0.95$, $p<.001$; $F2(2, 67) = 29.30$, $MSE = 0.59$, $p<.001$]. These results suggest that closed penults, palatal final onsets, and CV.CV.CV control items were treated differently. Proparoxytones and paroxytones behaved differently as well. Crucially, however, the significant interaction indicates that the patterning of the three word types differed by stress placement.

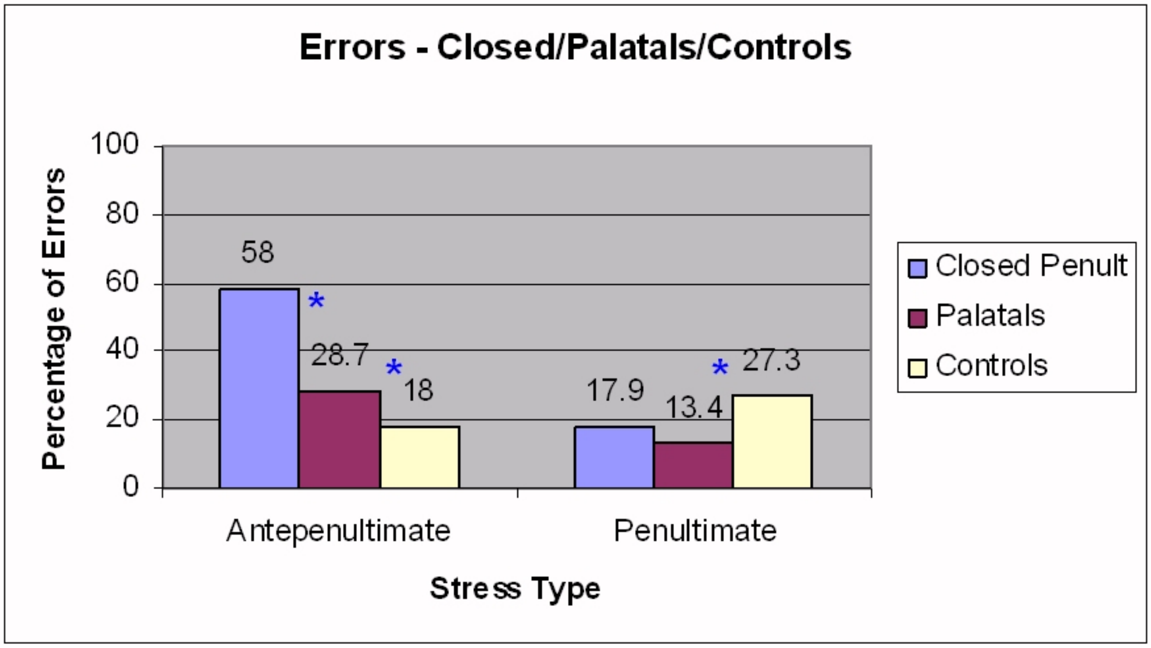

Figure **13**: Experiment 1: Error Rates – Closed/Palatals/Controls

As seen in figure 13, in the antepenultimate condition the palatal stimuli pattern between the proscribed closed penult stimuli and the control gaps (all pairwise comparisons in the antepenultimate condition = *p*<.01). Traditional accounts of syllable weight which call for a binary division between heavy and light syllables are unable to account for the tertiary division of error rates illustrated above. These data suggest that native speakers of Spanish are able to track the statistical patterns of phonotactic combinations in their lexicon at a much subtler level of generalization than accounted for by previous descriptive analyses. That is to say, while it is true that speakers differentiate between the proscribed cv́.cvc.cv pattern and the permitted cv́.cv.cv pattern, they are also able to cue into the integration of subsyllabic phonotactics (such as onsets) and higher-level prosodic patterns such as stress placement. The data in Experiment 1 suggest that speakers identify patterns such as cv́.cvc.cv as proscribed, as evidenced by very high error rates. They identify patterns such as cv́.cv.cv as generally acceptable. However, cv́.cv.pv patterns (where p = palatal) are identified as unattested in the lexicon. This distinction between attested patterns, unattested patterns, and proscribed patterns is borne out in the three-way split of error rates in the antepenultimate condition.

The behavior of the three word types in the penultimate condition corroborates this finding. In this condition we find no significant difference between the unattested palatal pattern and the proscribed closed penult category ($p>.66$). No significant differences among the three word types might appear to offer the strongest support for gradience in the phonotactic acceptability of native Spanish speakers. However, the higher level of errors for the control stimuli is arguably accounted for due the fact that the forms in this category are the only ones that are licit in either the antepenultimate or the penultimate conditions. Due to the large number of proparoxytone stimuli in the experiment in general, a higher error rate for the controls may be expected. It is noteworthy, however, that the unattested palatal final onsets and the proscribed closed penult patterns provoke substantially lower error rates in the penultimate condition ($p<.03$), as penultimate stress is predicted for heavy syllables.

Reaction time data for these three categories were also submitted to 2 (stress type) x 3 (word type) repeated measures ANOVAs by participants and by items. 5 participants were excluded from this analysis due to empty cells. The results found no significant main effect of stress type [$F1(1, 25) = 2.91$, $MSE = 17926$, $p>.10$; $F2(1, 67) = 0.07$, $MSE = 229$, $p>.80$], a main effect of word type that was significant only by participants [$F1(2, 50) = 11.38$, $MSE = 46690$, $p<.001$; $F2(2, 67) = 2.82$, $MSE = 16406$, $p>.06$], and an interaction of stress type and word type that was significant only by participants as well [$F1(2, 50) = 8.59$, $MSE = 36199$, $p<.001$; $F2(2, 67) = 2.77$, $MSE = 9807$, $p>.07$]. These findings suggest that the participants did not react differently to proparoxytones and paroxytones, although the word types did behave differently. The following figure illustrates the reaction times for the three word types in both stress conditions.

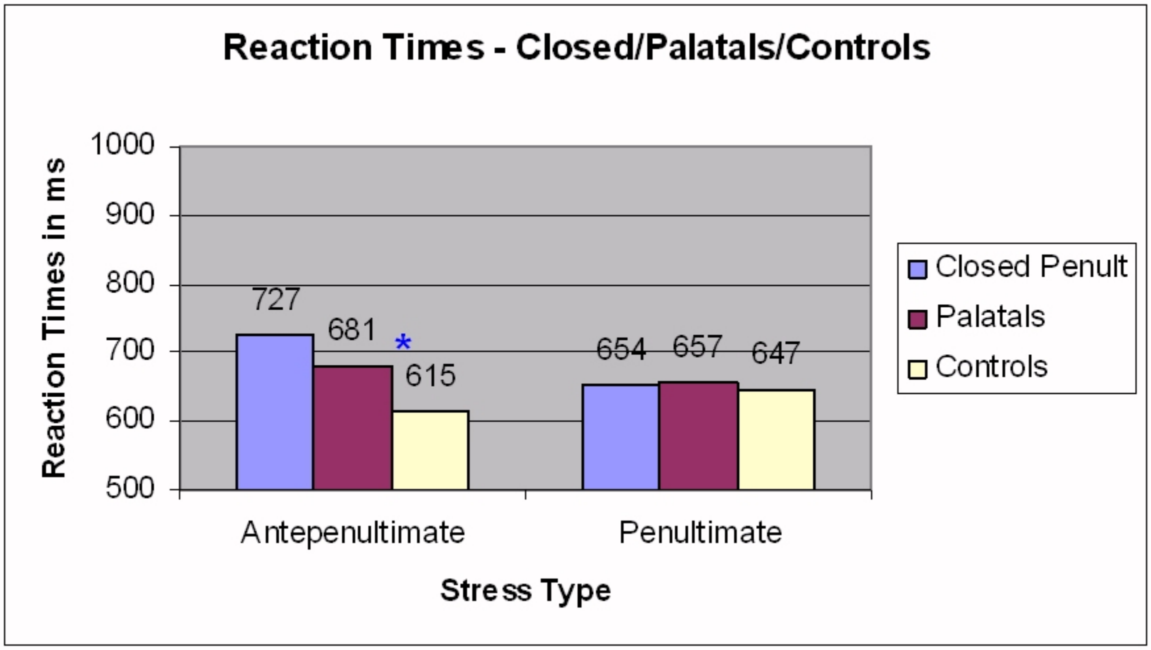

Figure **14**: Experiment 1: Reaction Times – Closed/Palatals/Controls

As in the previous latency analysis, the reaction times for the closed penults, palatals and controls partially corroborate the findings from the accuracy data. As predicted theoretically, there are no significant differences between the categories in the penultimate condition (p>.99). In the antepenultimate condition, post-hoc tests reveal a significant difference between palatals and controls (p<.01), but no difference between the palatals and the closed penults (p>.12). Again, the latency analysis was reduced in statistical power, given that only 84% of the data were available after inaccurate trials were eliminated. Nonetheless, the results still stand in contrast to traditional accounts. Palatal final onsets should pattern with the control items theoretically. However, these data suggest they are different. Instead, the palatal final onsets pattern with the proscribed forms.

The three-way split evidenced by the palatal final onset stimuli is an important finding in that it suggests that modern speakers of Spanish are differentially sensitive to both that which is deemed to be theoretically prohibited by the synchronic grammar as well as to structures that are absent for historical reasons. This is of particular relevance for models of synchronic phonology in that the theoretical architecture of such models needs to take into account a level of granularity that is usually not represented in, for example, approaches to Spanish stress which adduce a binary light/heavy syllable-structure distinction to account for restrictions on the trisyllabic stress window in the language. In such models, forms that represent phonotactic gaps hold no special status, and as such, are arguably not predicted to pattern distinctly from other phonologically licit but unattested items. Under an emergentist or usage-based view, however, speakers are predicted to be sensitive to a range of statistical patterns (at multiple levels of generalization) across the lexicon. The data from Experiment 1 suggest that, while speakers do maintain a phonological grammar that globally proscribes patterns such as proparoxytone forms with a closed penult, they also encode lower-level phonotactic knowledge that is reflected in their production of novel forms. Thus, while there may be no strict prohibition for Spanish speakers against proparoxytones of the shape CV.CV.PV, they are nevertheless highly attuned to the fact that the statistical probability of encountering such forms is zero. In this sense, diachrony persists in the synchronic grammar of Spanish in the form of statistical patterns (or the lack thereof) across the lexicon.

2.6.3 Rising and falling diphthongs

As discussed in chapter 1, traditional accounts of syllable weight characterize both rising and falling diphthongs as equally *heavy*. Both diphthong types have been shown to reduce the Spanish three-syllable window to one of two syllables. However, rising diphthongs do not always pattern with falling diphthongs with regards to attracting stress. Whereas falling diphthongs systematically attract stress, rising diphthongs do not (see §1.3.4). In the discussion of the general analysis of this chapter, rising and falling diphthongs were combined into one category to allow for comparison between various types of phonotactic constraints. However, it has been shown in section 2.6.2 that native speakers are sensitive to patterns that are more finely-grained than traditional theories have been able to account for. The participants in this experiment may also be sensitive to the differential behavior of the rising diphthongs across the Spanish lexicon. In a naming task, traditional theories of syllable weight would predict that either diphthong will provoke significantly more production errors and/or be articulated more slowly than comparable monophthong controls. However, sensitivity to the differences between the rising and falling diphthongs may lead participants to react distinctly when presented with nonce proparoxytones with a rising vs. falling diphthong in the penultimate syllable. In order to test this hypothesis, a separate ANOVA comparing rising and falling diphthongs in both the antepenultimate and penultimate stress conditions was performed on accuracy and latency data.

2 (stress type) x 2 (diphthong type) repeated measures ANOVAs by participants and items comparing the error rates of both diphthong types found a significant main effect of stress type [$F1(1, 30) = 73.22$, $MSE = 4.48$, $p<.001$; $F2(1, 26) = 136.20$, $MSE = 2.06$, $p<.001$], a significant main effect of diphthong type [$F1(1, 30) = 5.13$, $MSE = 0.22$, $p=.03$; $F2(1, 26) = 4.42$, $MSE = 0.10$, $p<.05$], and no significant interaction of stress type and diphthong type [$F1(1, 30) = 0.47$, $MSE = 0.00$, $p>.83$; $F2(1, 26) = 0.06$, $MSE = 0.00$, $p>.80$]. Participants treated proparoxytones and paroxytones differently. They also differentiated between the rising and falling diphthongs. However, the results of this analysis do not show that the patterning of the two diphthong types modulated stress placement. The following figure illustrates the error rates for rising and falling diphthongs in this experiment.

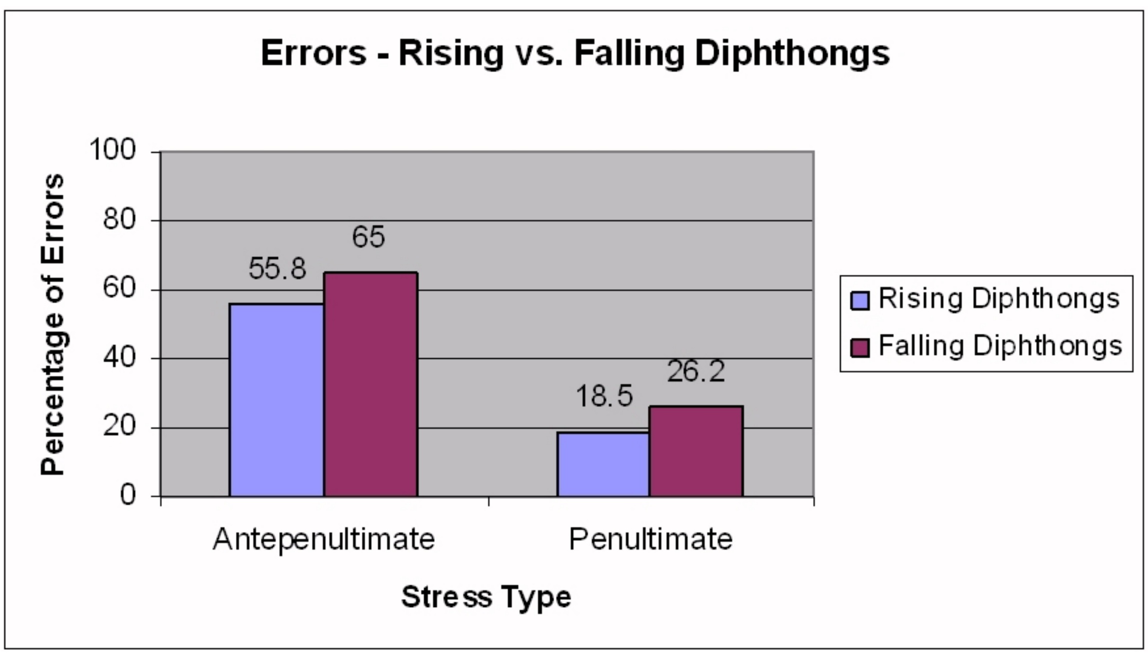

Figure **15**: Experiment 1: Error Rates – Rising vs. Falling Diphthongs

Descriptively, the results illustrated in figure 15, portray exactly what we predicted. The falling diphthongs provoked more error rates in the antepenultimate condition than the rising diphthongs. If the rising diphthongs are viewed as more "word-like", given their patterns within the lexicon, one would expect a substantial amount of errors in the antepenultimate condition but significantly fewer than for falling diphthongs, which systematically attract stress. This is the pattern evidenced in the data, but post-hoc analyses show that the difference between the diphthong types is not significant in either stress condition.

Reaction time data for the two diphthong types were also submitted to analyses of variance. However, 12 participants had to be excluded from the analyses due to empty cells. 2 (stress type) x 2 (diphthong type) repeated measures ANOVAs by participants and items comparing latency data found a main effect of stress type that was significant by participants [$F1(1, 18) = 6.97$, $MSE = 68299$, $p<.02$; $F2(1, 26) = 0.02$, $MSE = 115$, $p>.90$], no significant effect of diphthong type [$F1(1, 18) = 0.65$, $MSE = 3743$, $p>.43$; $F2(1, 26) = 0.02$, $MSE = 108$, $p>.89$], and no significant interaction of stress type and word type [$F1(1, 18) = 1.49$, $MSE = 6350$, $p>.23$; $F2(1, 26) = 0.04$, $MSE = 280$, $p>.84$]. Again, participants reacted different to proparoxytone and paroxytone forms, but there was no differential treatment of the diphthong types. The follow figure depicts the latency data for the diphthong types across stress conditions.

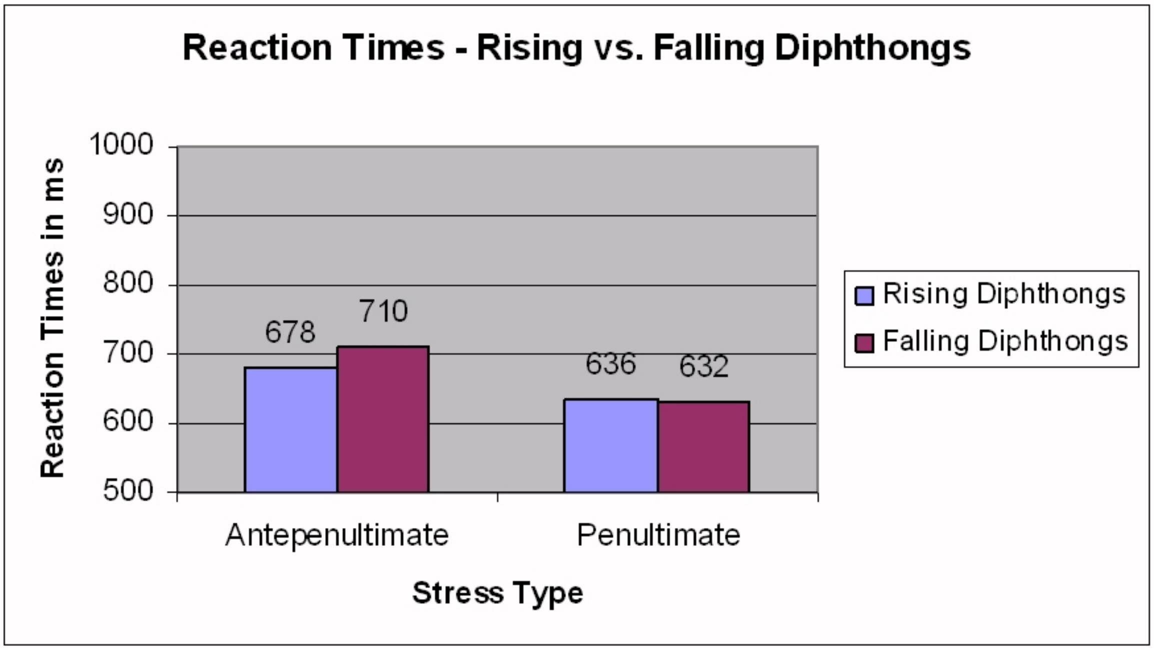

Figure **16**: Experiment 1: Reaction Times – Rising vs. Falling Diphthongs

While descriptively it appears that the falling diphthong were pronounced more slowly than their rising diphthong counterparts, this difference is not significant (p>.44). The elimination of almost 40% of the latency data and the low quantity of items of each diphthong type appear to prevent the predicted diphthong effect from surfacing in this analysis as well.

The data presented in figures 15 and 16 suggest that the pattern predicted by our observations may be present, but the statistical power of the analysis is not sufficiently strong to support any claims. There may be too many other variables included in this first experiment for the diphthong patterns to surface as statistically significant. Given the fact that there were five total phonotactic constraints at play in this experiment, in addition to 140 real word filler items, it is understandable that the subtler effects may need more items in order to emerge. Chapters 3 and 4 examine the rising vs. falling diphthong effect more closely by increasing the amount of diphthong stimuli and eliminating the other phonotactic constraint categories.

2.6.4 Fillers

Finally, reaction time data for real word filler items were also submitted to 2 (frequency) x 3 (stress type) ANOVAs. These analyses were performed to test the validity of the naming task. One may argue that the task of naming nonwords is artificial. A counterargument to the findings discussed in this dissertation could be that the participants engaged reading strategies that are substantively different from their habitual reading for real words. One way to verify that the participants are reading naturally, both for real words and nonwords, is to confirm the replication of a recognized effect from the literature. It is well established that speakers produce high frequency words more quickly than low frequency words (Oldfield & Wingfield, 1965; Jescheniak & Levelt, 1994; Griffen & Bock, 1998; Caramazza et al., 2001; Schatzman & Schiller, 2004). Therefore, if the frequency effect is found in the real word fillers used in Experiment 2, we can reasonably conclude that participants attended to the task and that the naming task served its function. Results found a significant main effect of stress type [$F1(2, 60) = 9.28$, $MSE = 5549$, $p<.001$; $F2(2, 134) = 4.99$, $MSE = 4481$, $p<.01$], a significant main effect of frequency [$F1(1, 30) = 50.42$, $MSE = 54410$, $p<.001$; $F2(1, 134) = 46.07$, $MSE = 41387$, $p<.001$], and an interaction of stress type and frequency that was significant by participants [$F1(2, 60) = 5.14$, $MSE = 1640$, $p<.01$; $F2(2, 134) = 1.55$, $MSE = 1391$, $p>.21$]. These data evidence a clear frequency effect. The following figure illustrates the reaction times for high and low frequency real-word fillers across stress conditions.

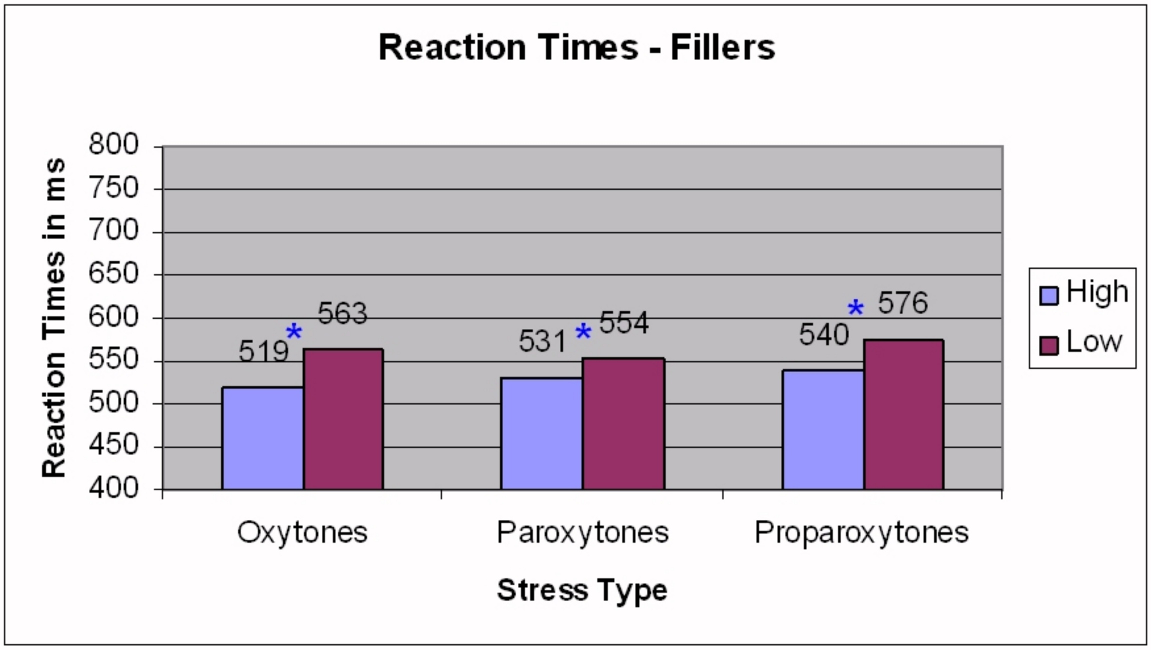

Figure **17**: Experiment 1: Reaction Times – Fillers

High frequency fillers were produced significantly more quickly in all stress conditions ($p<.001$).

2.6.5 General discussion

The findings in Experiment 1 validate the naming task as an appropriate tool to probe native speakers' knowledge of phonotactic constraints in their language. It is clear that the adaptation of methodology from cognitive psychology opens up new possibilities to the linguist who is interested in the cognitive representation of linguistic structure. The results discussed in this chapter both confirm and challenge traditional theories of syllable weight and Spanish phonotactics. We know the naming task serves its function in that the theoretical weight distinction between *heavy* and *light* syllables is borne out in data. Syllables traditionally characterized as *heavy* do provoke higher error rates than *light* syllables. This distinction appears only in the stress condition where a phonological constraint on stress placement predicts it should occur. This finding both submits empirical evidence consequent with weight sensitive theories of Spanish stress and validates the use of behavioral methods to test such theories.

The methodology employed in Experiment 1 offers a novel approach to test well-established patterns in the literature in that it is able to tune into far more finely-grained patterns without resorting to intuition and/or metalinguistic reflection. This can be seen in the analysis of words whose final syllables contain palatal consonant onsets. Abstract, symbolic approaches to Spanish syllable structure and stress placement cannot easily account for both proscriptions deriving from phonological restrictions on stress assignment in the synchronic grammar and the sensitivity attested here to diachronic gaps across the lexicon. Experiment 1 shows that speakers are attuned to both levels of generalization.

While not supported statistically, the descriptive analysis of the rising and falling diphthong stimuli in this experiment suggest that a subtler pattern between two forms that are theoretically equally *heavy* may exist. Given the overall design of Experiment 1, this pattern did not surface fully. However, an experimental design which focuses more closely on the diphthongs themselves might render interesting results in favor of further gradience within the system. This is the goal of Experiment 2.

Chapter 3

Experiment 2: Speeded Naming Task - Diphthongs

3.1 Introduction

Experiment 2 examines more closely the interaction of rising and falling diphthongs with stress placement in Spanish. As discussed in chapter 1, both rising and falling diphthongs constitute heavy syllables within traditional approaches to Spanish syllable structure. Given the binary (light versus heavy) nature of such accounts of syllable weight, traditional phonological theory does not predict any difference in behavior between the diphthong types. Specifically, for our purposes here, when participants are presented with penultimate stress in a naming task, both rising and falling diphthongs in the penultimate syllable should attract stress and provoke few errors. And if diphthongs attract stress to the penultimate syllable, antepenultimate stress should be banned. Therefore, if a word form with a diphthong in the penultimate syllable is marked for antepenultimate stress, the prediction is that participants will commit many more production errors, take longer to produce the forms, or a combination of both.

Despite the fact that traditional theory does not yield any predictions regarding the difference in behavior between the rising and falling diphthongs, there is descriptive evidence that the two diphthong types behave differently. As elaborated in chapter 1, rising diphthongs do not systematically attract stress in the final syllable, whereas falling diphthongs always attract stress in this position. Therefore, a statistical approach to the lexicon predicts that native speakers may be sensitive at a more general level to the different patterning of the rising diphthongs with respect to stress placement. If this is true, we might predict differential treatment of the two diphthong types in an appropriately designed behavioral task.

With this goal in mind, Experiment 2 tests the way monolingual speakers of Spanish process rising and falling diphthongs, and more specifically, how the different diphthong types modulate stress assignment. This analysis is similar to Experiment 1 in that it tests the nature of varying phonotactic constraints, but differs in that it is a focused study of the diphthong patterns.

3.2 Participants

23 functionally monolingual adult speakers of Spanish participated in Experiment 2. All were undergraduate students at the University of Jaen, Spain. One participant's data were excluded due to technical errors with digital recording.

3.3 Materials

As this second experiment focuses specifically on how the different diphthong types modulate stress placement in Spanish, the diphthong stimuli developed for the previous experiment were expanded to include 56 nonword diphthong stimuli. Critical stimuli were comprised of three-syllable words with a diphthong in the penultimate syllable. The diphthongs in the penultimate syllable consisted of either a rising diphthong (CV.C**GV**.CV) or a falling diphthong (CV.C**VG**.CV). Rising diphthongs were represented by either *ia* or *ie* (e.g. *dóbiana* or *góviego*), and falling diphthongs were represented by *ai* or *ei* (*mómaino* or *fáteiga*), orthographically. The segmental makeup of these stimuli was strictly controlled so that all consonants and vowels in the antepenultimate and final syllables appeared equally across all conditions.

A set of monophthong control stimuli was also designed to serve as a comparison to the diphthong condition. These stimuli matched the diphthong stimuli segmentally, with the exception of the penultimate syllable, which contained a monophthong vowel instead of the diphthong (e.g. *fátaga*).

All diphthong and monophthong stimuli were presented both with and without a diacritic accent mark over the antepenultimate syllable. This allowed for the comparison of segmentally identical stimuli with either proparoxytone or paroxytone stress. However, to avoid any possible priming effects that might occur if participants saw segmentally identical stimuli in both conditions, stimuli with and without accent marks were presented in a counter-balanced fashion. That is to say, if participant 1 saw *dóbiana*, then participant 2 saw *dobiana*, and so forth. These stimuli total 112 critical items seen by each participant.

In addition to the critical items, participants also saw 168 filler stimuli. Nonword stimuli consisted of 28 CV.CV.CV items that were not controlled to be segmentally similar to the critical items. Additionally, 140 real words were chosen from the Alameda & Cuetos (1995) database. These words consisted of both high and low frequency items in three categories of antepenultimate, penultimate and final stress. This rendered a total of 140 total nonword items and 140 total real words. A complete list of stimuli for this experiment is listed in Appendix B.

The following table lists examples of the critical items and nonword controls in this experiment with quantities listed in parenthesis.

Table **9**: Stimuli summary for Experiment 2

Critical Items:	*Antepenultimate*	*Penultimate*
Rising Diphthongs	dóbiana (28)	goviano (28)
Falling Diphthongs	fáteiga (28)	zateino (28)
Monophthong Controls	mósago (56)	morega (56)

The following diagram exemplifies what each participant saw over the course of the experiment.

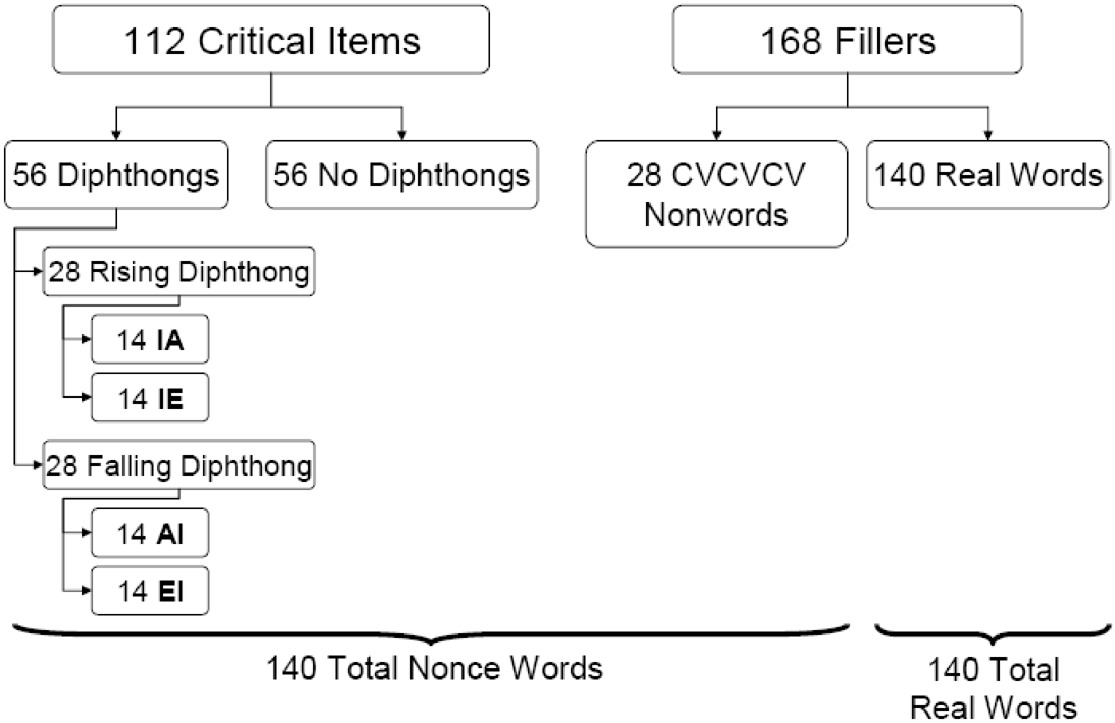

Figure **18**: Visual overview of Experiment 2

3.4 Procedure

Participants were tested individually in a quiet room. They were seated in front of a computer screen and two microphones. The first microphone was connected to a PST Serial Response Box, and the second to a digital recorder (Marantz PMD660 sampling at 48 kHz). The button box served as a voice key to cue the computer to record reaction times and to present the next stimulus when the participant spoke. The digital recorder was used to record the experimental session for subsequent coding of the accuracy of the participants' responses.

In this experiment the participants performed a standard naming task. First, an instructions screen (in Spanish) informed the participants that they would be reading words one by one on a computer screen, and that they should read the words aloud as quickly and accurately as possible. Following the instructions screen, participants performed a practice session with 12 items to become accustomed to the task.

Each trial began with a fixation point (+) which was presented in the middle of the screen in black font on a white background. This was done to focus the participants' attention to the center of the screen and to cue them that the word they needed to pronounce was about to appear. The fixation point remained on the screen for 500 milliseconds and was followed by a blank screen for 500 milliseconds. Following these 1000 milliseconds, the stimulus was presented and remained on the screen until the participant spoke. The participants' speech cued the voice key, which caused the stimulus on the screen to disappear and be replaced with a blank white screen. During the blank screen, the experimenter recorded on the button box the stress the participant assigned to the stimulus (i.e. antepenultimate, penultimate, or final). In the case that the participant committed an error, the experimenter cued the computer to present an error screen. The word "Error" appeared in red letters in the upper left-hand corner of the screen. The purpose of this error message was to alert the participant to the fact that they had committed an error and to focus them on the task. After the error message, another blank screen appeared for 1000 milliseconds until the following trial began.

Once the participants completed the practice trials, the experimental stimuli were presented in random order over two experimental blocks with a 30 second break in between them. The following diagram illustrates the standard naming task as it was used in this experiment.

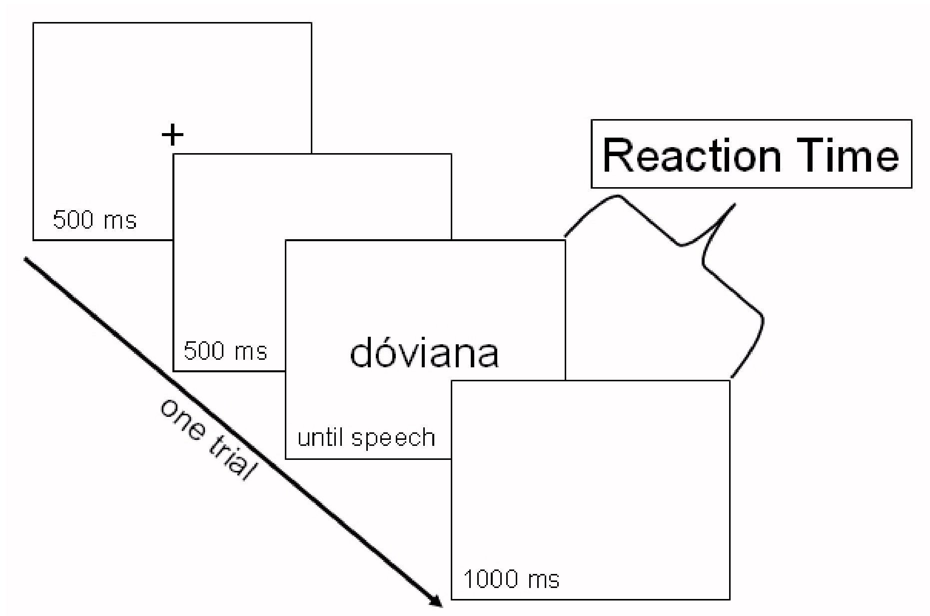

Figure **19**: Procedure for Experiment 2 (standard naming task)

3.5 Data Analyses

Both error rate analyses and reaction time analyses were conducted on the data from this experiment. Native or near-native listeners (including the author) coded the errors manually using digital recordings of the experimental sessions. Errors comprised two categories: production errors and technical errors. Production errors consisted of instances in which participants misspoke the stimulus. Examples of this type of error were stress shifts to another syllable, changes in the segmental makeup of the word, or disfluencies in which words were not finished or contained pauses. Technical errors were those in which the microphone or the response box failed. These types of errors resulted in either extremely long reaction times, such as when the microphone did not trigger the voice key after articulation, or extremely fast reaction times, such as when the voice key triggered the next stimulus prematurely before the participant actually spoke.

The data were cleaned as follows. First, technical errors due to the microphone and/or voice key were eliminated from all analyses. Production errors were included in error rate analyses, and excluded from reaction time analyses. All reaction time latencies were cleaned to exclude any responses faster than 200 ms and slower than 2000 ms. Subsequently, any latencies 2.5 standard deviations above or below each individual's mean reaction time were also eliminated.

Both reaction time and error data were submitted to separate repeated measures ANOVAs by participants (*F1*) to test for variance across the population and by items (*F2*) to test for variance across the stimuli set. Main effects and interactions were followed up by Tukey's HSD post-hoc tests to identify any independent effects across the experimental categories and conditions. Results are discussed below.

3.6 Results and Discussion

ANOVAs by participants and items of error rates comparing diphthongs to their monophthongs controls found a significant main effect of word type (diphthong vs. monophthongs in the penult) [$F1(1, 21) = 26.41$, $MSE = 0.26$, $p<.001$; $F2(1, 105) = 25.08$, $MSE = 0.64$, $p<.001$], a significant main effect of stress type (proparoxytone vs. paroxytone) [$F1(1, 21) = 12.55$, $MSE = 0.21$, $p<.002$; $F2(1, 105) = 22.79$, $MSE = 0.53$, $p<.001$], and a significant interaction of word type and stress type [$F1(1, 21) = 69.09$, $MSE = 0.93$, $p<.001$; $F2(1, 105) = 95.81$, $MSE = 2.22$, $p<.001$]. The results clearly show that participants treat diphthong and monophthongs stimuli differently. They also react distinctly to words in the paroxytone and proparoxytone conditions. Crucially, the results also find that participants treat the diphthong and monophthongs stimuli differently in each stress condition. The following figure displays the error rates for the diphthongs and the monophthongs controls in both the antepenultimate and penultimate conditions.

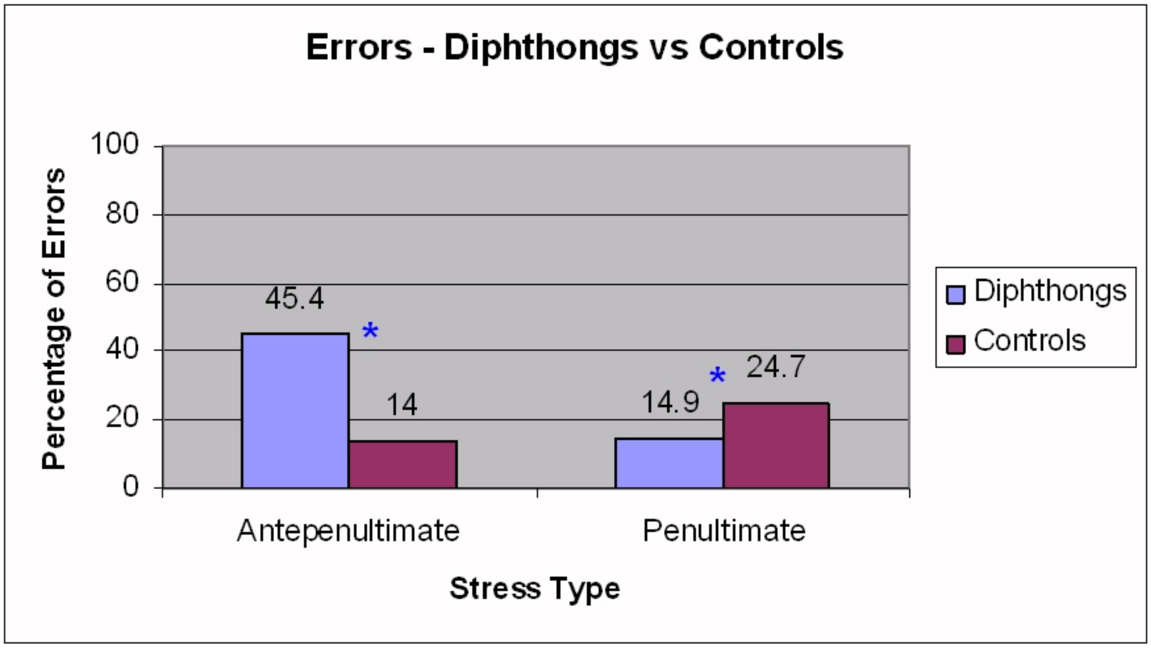

Figure **20**: Experiment 2: Error Rates - Diphthongs vs. Controls

The antepenultimate condition reflects those stimuli that were presented with a diacritic over the first syllable to cue proparoxytone stress. If diphthongs in the penultimate syllable attract stress, then forcing participants to stress the antepenultimate syllable should provoke a high level of errors. Monophthongs in this condition are phonologically licit. Therefore, the prediction is that control items will provoke significantly fewer errors when presented with proparoxytone stress. This is clearly the finding evidenced by Experiment 2. In the antepenultimate condition, participants committed 31.4% more errors when presented with diphthong stimuli (p<.001).

This finding offers robust support for traditional theories of syllable weight that call for a division between heavy and light syllables. Diphthongs form heavy syllables which attract stress. They also, in turn, inhibit stress assignment to any syllable further to the left, when assigning stress from the right edge of the word (as is the case in Spanish). Light syllables do not attract stress. Therefore, stress assignment to syllables further to the left is not prohibited by the grammar. Experiment 2 evidences this exact pattern.

Arguably, these arguments are only accurate if the pattern is not exhibited in the penultimate condition as well. That is, if the diphthong stimuli also provoke significantly more errors in the penultimate condition, then it could be argued that under the demands of the naming task, diphthongs are simply more difficult to pronounce in general, regardless of the their putative phonological weight. The results in figure 20, however, evidence the opposite pattern in the penultimate condition. In this condition the monophthong controls provoked significantly more errors than the diphthong stimuli ($p<.05$). Considering that the penultimate condition falls clearly within the two-syllable window, and therefore should license any syllable structure in the penultimate syllable, one might expect no significant difference between either word type in this condition. Such a prediction falls clearly in line with the weight sensitive account of the Spanish stress system. However, the findings here are also easily explained within the same account. The higher percentage of errors among the control items can be attributed to the fact that CV.CV.CV structures are licit forms in either the antepenultimate or penultimate conditions. Therefore, a certain level of variance is expected. The high level of proparoxytones in the experiment overall is likely to influence the error rates in the control condition. At the same time, a weight sensitive account of Spanish stress argues that *heavy* syllables will attract stress. Given that the diphthongs lie in the penultimate syllable of the experimental items, the lower level of errors among the diphthong stimuli in the penultimate condition is indicative of the fact that they attract stress to that position.

In this sense, the results of Experiment 2 accord with accounts that call for a quantity sensitive interpretation of Spanish stress (Harris, 1983, 1992; Lipski, 1997). However, Experiment 2 had two goals. The second purpose of Experiment 2 was to contrast the behavior of the two diphthong types, given the more ambiguous overall patterning of the rising diphthongs in Spanish with respect to the putative weight question. Specifically, we hypothesized that participants might treat rising diphthongs as "more word-like" than falling diphthongs, given their less globally consistent pattern of attracting stress. To test this hypothesis, additional 2 (diphthong type) x 2 (stress type) ANOVAs were run.

The second analyses of variance by participants and items of error rates found a significant main effect of diphthong type [$F1(1, 21) = 14.67$, $MSE = 0.25$, $p=.001$; $F2(1, 50) = 11.33$, $MSE = 0.29$, $p<.002$], a significant main effect of stress [$F1(1, 21) = 105.25$, $MSE = 2.06$, $p<.001$; $F2(1, 50) = 162.40$, $MSE = 2.39$, $p<.001$], and a significant interaction of diphthong type and stress type [$F1(1, 21) = 17.02$, $MSE = 0.18$, $p<.001$; $F2(1, 50) = 13.84$, $MSE = 0.20$, $p<.001$]. These results show that participants do indeed treat rising and falling diphthongs differently. And most importantly, the behavior of the two diphthong types varies according to stress type as well. The following figure illustrates the percentage of errors produced by participants when presented with either rising or falling diphthongs in both the antepenultimate and penultimate conditions.

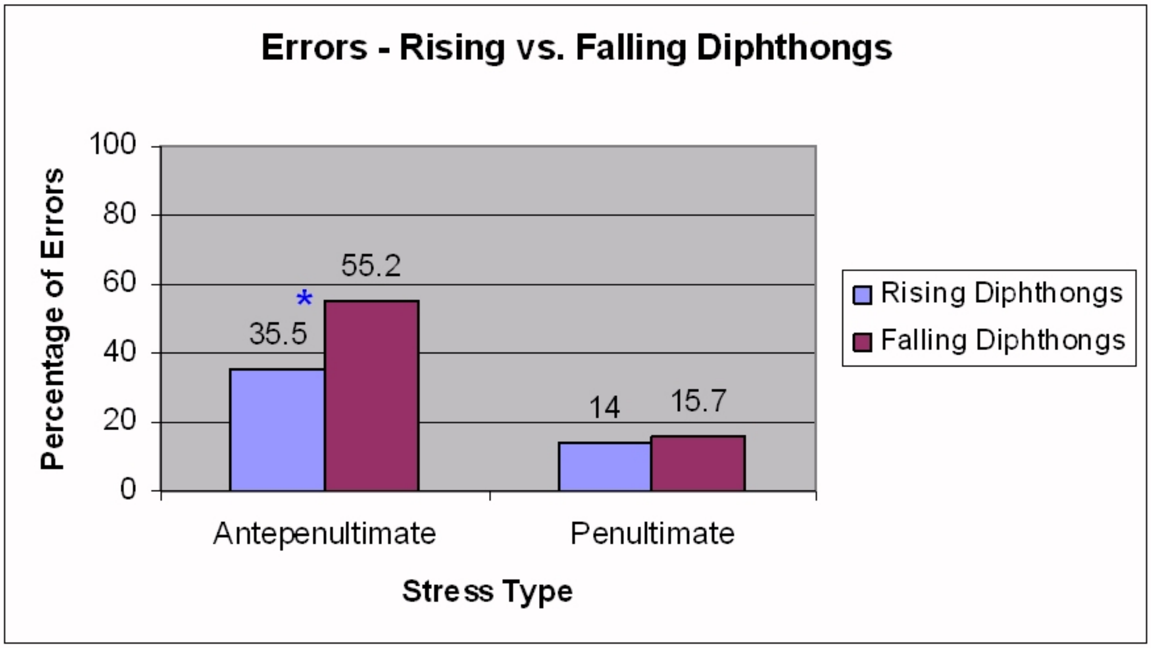

Figure **21**: Experiment 2: Error Rates – Rising Diphthongs vs. Falling Diphthongs

The crucial finding demonstrated in figure 21 is found in the antepenultimate condition. These results suggest that native speakers of Spanish cognitively represent and process rising and falling diphthongs differently. This is evidenced by the fact that falling diphthongs, those forms predicted to represent *heavier* syllables, provoke higher error rates than the rising diphthongs ($p<.001$).

In the penultimate condition, we find no significant difference between diphthong types. This is robust evidence in favor of the three-syllable window, given that the subtle finding between rising and falling diphthongs only surfaces in a condition where stress assignment is predicted to be phonologically proscribed. Since both rising and falling diphthongs attract stress to the penultimate syllable, 1) a much lower level of error is evidenced in the penultimate condition, and 2) there is no significant difference between the diphthong types (p>.94).

Reaction time data were also submitted to 2 x 2 repeated measures ANOVAs. Analyses comparing diphthongs to their monophthongs controls reveal a significant main effect of word type [$F1(1, 21) = 14.84$, $MSE = 60182$, p<.001; $F2(1, 105) = 6.55$, $MSE = 55841$, p<.02], a significant main effect of stress type [$F1(1, 21) = 18.33$, $MSE = 23710$, p<. 001; $F2(1, 105) = 4.91$, $MSE = 21004$, p<.03], and a significant interaction of word type and stress type [$F1(1, 21) = 10.30$, $MSE = 17093$, p<.01; $F2(1, 105) = 5.89$, $MSE = 25207$, p<.02]. These results support the error data that indicate that participants treat words with diphthongs differently than words with monophthongs in the penult. They also corroborate the previous finding that participants behave differently when presented with words marked for proparoxytone vs. paroxytone stress. Most importantly, again, we find a significant interaction which is indicative of the fact that the distinct word types modulate stress placement.

The following figure illustrates the reaction time data for the diphthong stimuli and their controls when presented with both antepenultimate and penultimate stress.

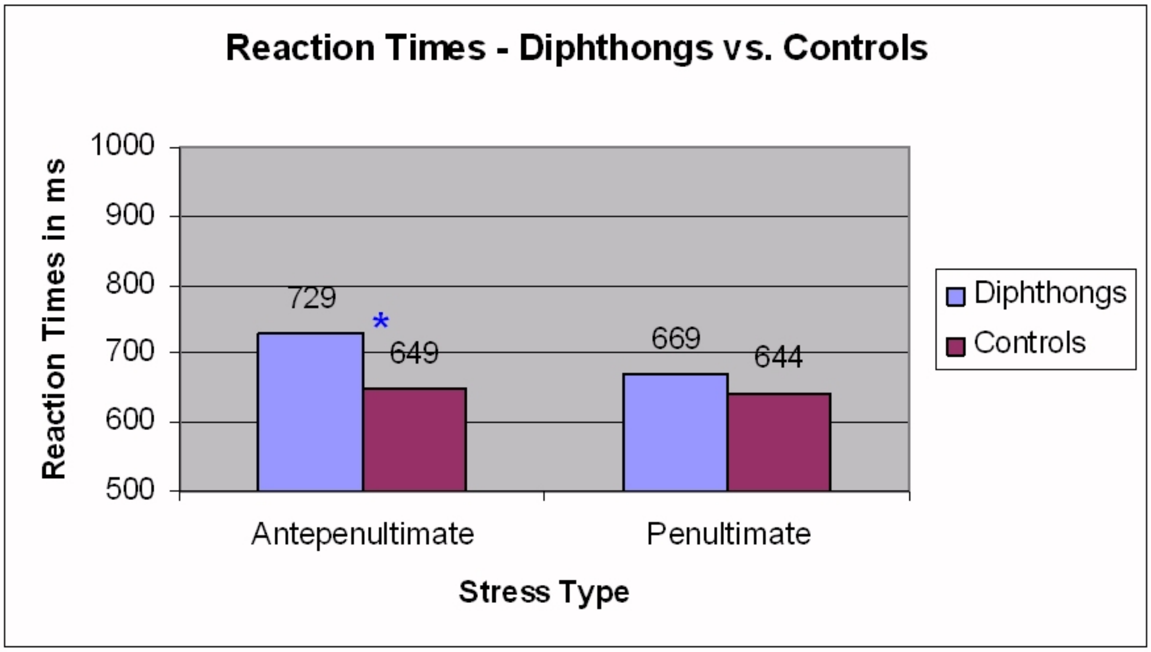

Figure **22**: Experiment 2: Reaction Times – Diphthongs vs. Controls

As illustrated in Figure 22, the diphthong stimuli yielded significantly longer naming latencies than monophthong controls when presented in the proparoxytone stress condition (p<.001). The difference between word types is not significant in the penultimate condition (p>.22). This is indicative of the fact that diphthongs in the penultimate syllable of a proparoxytone word are proscribed by the phonological grammar. In contrast, the monophthongs controls are phonologically permitted with either proparoxytone or paroxytone stress. The reaction time data reflect this fact in that there is no significant difference between participant reaction times for control stimuli across stress conditions (p>.97).

The analysis of reaction times to test the subtler difference between diphthong types is more difficult given the high error rates. The prohibition on proparoxytones with diphthongs in the penultimate syllable is so robust (as discussed previously) that it produces so many errors that reaction time data on correct trials is greatly reduced. In order to run ANOVAs by participants and items on reaction time data comparing diphthong type and stress type, one participant had to be eliminated due to empty cells. The results find a significant main effect of stress type [$F1(1, 20) = 20.15$, $MSE = 100777$, $p<.001$; $F2(1, 50) = 6.82$, $MSE = 44857$, $p<.02$]. However, no significant main effect of diphthong type was found [$F1(1, 20) = 0.01$, MSE = 30, $p>.90$; $F2(1, 50) = 1.35$, $MSE = 7795$, $p>.25$]. There was also no significant interaction of stress type and diphthong type [$F1(1, 20) = 0.23$, $MSE = 250$, $p>.63$; $F2(1, 50) = 0.67$, $MSE = 4427$, $p>.41$]. The main finding of this analysis is that participants read proparoxytones more slowly than paroxytones. This is expected, given that paroxytones represent the unmarked stress pattern in Spanish. However, the effect between diphthong types does not emerge in the reaction time data. The following figure represents the reaction time results for rising and falling diphthongs when presented with proparoxytone and paroxytone stress.

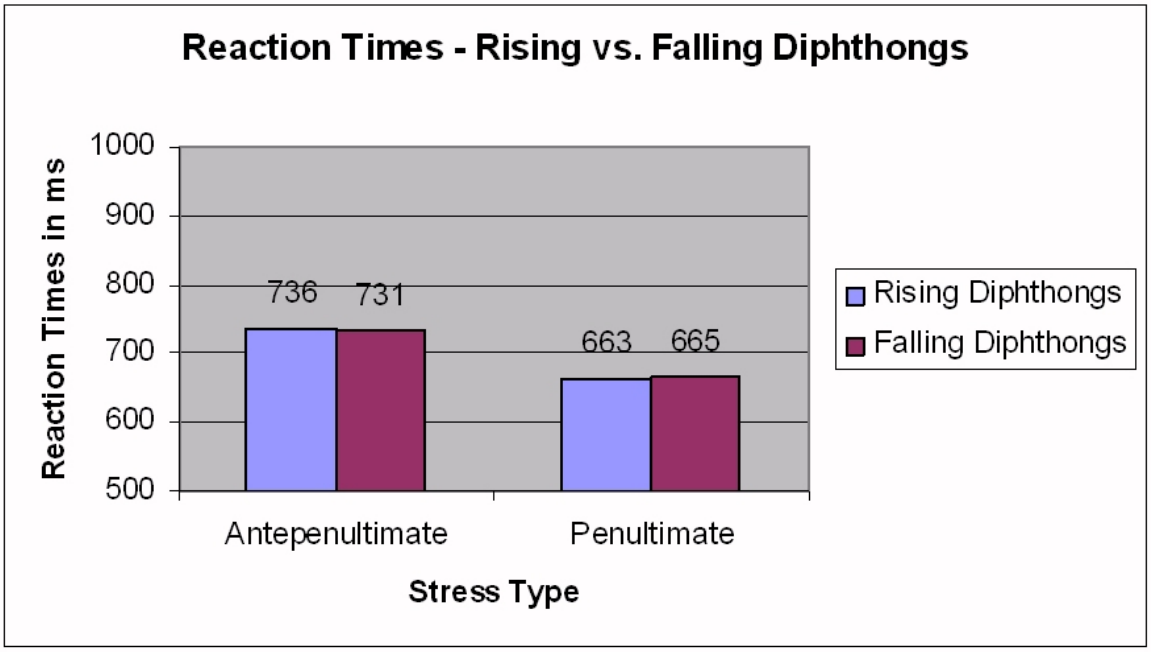

Figure **23**: Experiment 2: Reaction Times – Rising Diphthongs vs. Falling Diphthongs

As figure 23 illustrates, in general participants are slower to read proparoxytones than paroxytones. However, the difference between the rising and falling diphthongs is not significant in the reaction time data. The loss of the diphthong effect is arguably a result of high error rates. As illustrated in figure 21, over one third of all proparoxytones with rising diphthongs provoked production errors and over half of all proparoxytones with falling diphthongs were not pronounced accurately. This greatly reduced the amount of accurate trials from which reaction time data could be culled for analysis. Given the subtle nature of the differential pattern between diphthong types, the elimination of so many trials appears to wash away the effect. It is also possible that reaction time analyses are an inappropriate tool to test between two impossible phonotactic sequences. The error rate analyses appear to probe better the subtle differences between the diphthong types.

Lastly, data from real-word filler items were submitted to separate analyses of variance to test for the frequency effect. The ANOVAs for reaction time data on real word fillers found a significant main effect of stress type [$F1(2, 42) = 12.39$, $MSE = 10107$, $p<.001$; $F2(2, 134) = 10.72$, $MSE = 9354$, $p<.001$], a significant main effect of frequency [$F1(1, 21) = 21.07$, $MSE = 50049$, $p<.001$; $F2(1, 134) = 55.43$, $MSE = 48378$, $p<.001$], and no significant interaction of stress type and frequency [$F1(2, 42) = 2.69$, $MSE = 800$, $p>.07$; $F2(2, 134) = 0.78$, $MSE = 684$, $p>.45$]. As illustrated in the following figure, a clear frequency effect was obtained.

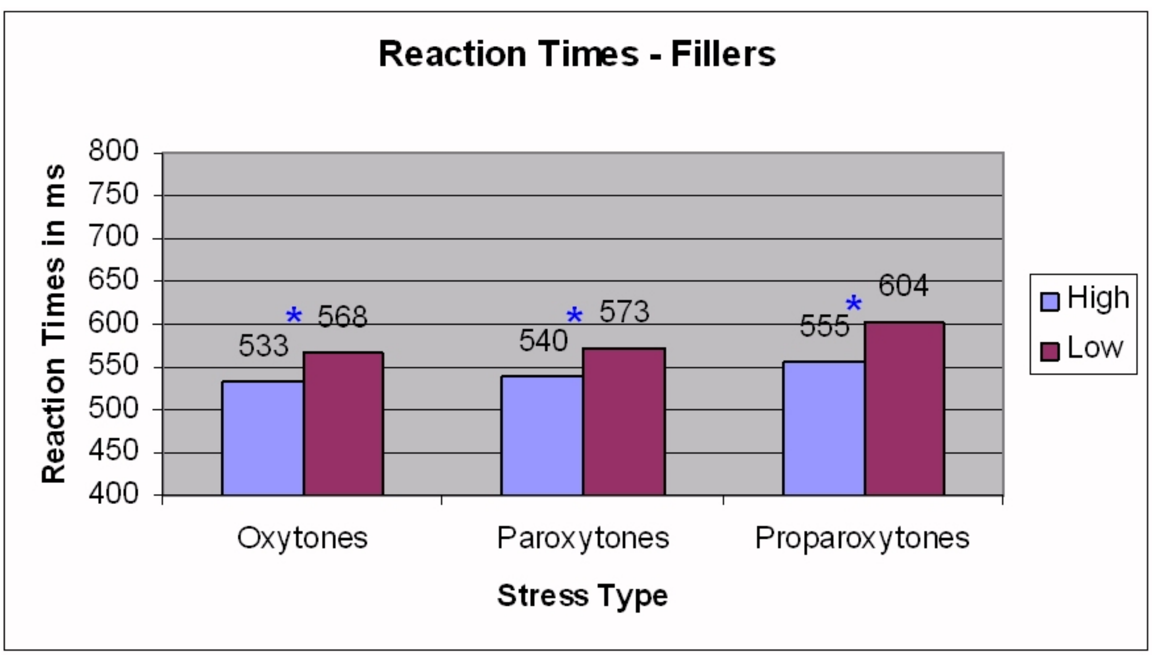

Figure **24**: Experiment 2: Reaction Times - Fillers

The main effect of stress type is due to the proparoxytone stimuli. Post-hoc analyses reveal a significant difference between proparoxytones and oxytones only in the high frequency condition ($p<.01$). Among the low frequency stimuli, there is a significant difference between the proparoxytone items and both paroxytones and oxytones ($p<.01$). Importantly, however, participants in this experiment were systematically slower to articulate low frequency real words as compared to high frequency words ($p<.001$).

In summary, these data reveal that, when presented in a highly sensitive behavioral task, the finely-grained differences between subtle phonotactic patterns may be tested empirically. These results shed new light onto issues of Spanish stress placement, and questions of syllable weight in general. It is clear that traditional accounts of a quantity sensitive interpretation of Spanish stress are supported in a broad sense by the results found in Experiment 2. There is a robust effect of word type in which diphthongs provoke far higher rates of errors than monophthongs controls. However, traditional representations fall short in their ability to account for more finely-grained patterns evidenced by this experiment. In contrast to the binary account of standard theories of syllable weight, the empirical data described above are more easily understood within frameworks arguing in favor of a conception of the lexicon in which the phonological grammar is built upon gradient constraints on syllable structure and stress assignment. A stochastic approach to the lexicon provides a much clearer context within which these results may be interpreted. If native speakers of Spanish are able to track the differential behavior of the diphthongs in the form of how they pattern across the Spanish lexicon, it becomes easier to understand how gradient restrictions on stress placement might arise.

Experiment 2 provides empirical data that support traditional theories of syllable weight and stress placement in Spanish in their interpretation of weight sensitivity (or weight modulated stress assignment) in Spanish. However, it goes further to find subtler patterns which traditional approaches cannot explain. The differential treatment of rising vs. falling diphthongs is so subtle, that it was not found in Experiment 1 where the quantity of diphthong stimuli was too low to drive the effect amongst other phonotactic constraints. Similarly, the elimination of inaccurate trials left too few accurate trials for the effect to surface in reaction time analyses in Experiment 2. Nonetheless, the error rate data in this experiment reveal a surprising, and robust, finding in which rising diphthongs pattern as *less heavy,* or *more word-like* in the context of nonword phonotactics, than do falling diphthongs.

Given the theoretical import of these findings, and the finely-grained nature of the effect, Experiment 3 seeks to replicate the findings discussed here with a different population. If we are able to find a similar pattern of results with Spanish speakers on another continent, this will offer strong empirical support for the argument that these findings do indeed represent a deep understanding of the Spanish lexicon and phonological system broadly shared by native speakers of Spanish.

Chapter 4

Experiment 3: Speeded Naming Task – Diphthongs Replication

4.1 Introduction

Experiment 3 tests how native speakers of a different dialect of Spanish process rising vs. falling diphthongs. Experiment 2 evidenced a significant difference between rising and falling diphthongs as to their modulation of stress placement. Specifically, rising diphthongs in the penultimate syllable appear to exhibit a weaker prohibition on proparoxytone stress than their falling diphthong counterparts. Both diphthong types were shown to provoke significantly more errors than the monophthong control stimuli, implying gradience in the system. This finding is unexpected, given that traditional theories of Spanish weight hold both diphthong types as equally *heavy*. Therefore, Experiment 3 explores this phenomenon with a different population. The purpose of experiment 3 is thus to replicate the findings of Experiment 2 and to do so with a different population of Spanish speakers. By testing a different population of speakers, Experiment 3 thus also tests for whether the results of Experiment 2 generalize to other parts of the Spanish speaking world.

4.2 Participants

36 functionally monolingual adult speakers of Spanish participated in Experiment 3. All were native Spanish speakers from Mexico and were undergraduate students at the Iberoamerican University in Puebla, Mexico.

4.3 Materials

The materials for this experiment are identical to those in Experiment 2.

4.4 Procedure

The procedure is also the same in Experiment 2 and Experiment 3.

4.5 Data Analyses

Error rates and reaction time data were analyzed as described in Experiment 2.

4.6 Results and Discussion

ANOVAs by participants and items of error rates comparing diphthongs to their monophthongs controls revealed a significant main effect of word type (diphthong vs. monophthongs in the penult) [$F1(1, 35) = 41.04$, $MSE = 0.40$, $p<.001$; $F2(1, 105) = 46.26$, $MSE = 0.62$, $p<.001$], a significant main effect of stress type (proparoxytone vs. paroxytone) [$F1(1, 35) = 52.37$, $MSE = 4.48$, $p<.001$; $F2(1, 105) = 317.70$, $MSE = 6.54$, $p<.001$], and a significant interaction of word type and stress type [$F1(1, 35) = 80.50$, $MSE = 1.67$, $p<.001$; $F2(1, 105) = 121.32$, $MSE = 2.50$, $p<.001$]. Participants reacted differently to diphthongs and monophthongs. They also treated words differently when presented with proparoxytone or paroxytone stress. Most importantly, participants' reactions to diphthongs and monophthongs were different for each word type in each stress condition. The following figure illustrates the percentage of error for diphthongs and their monophthong controls in each stress condition.

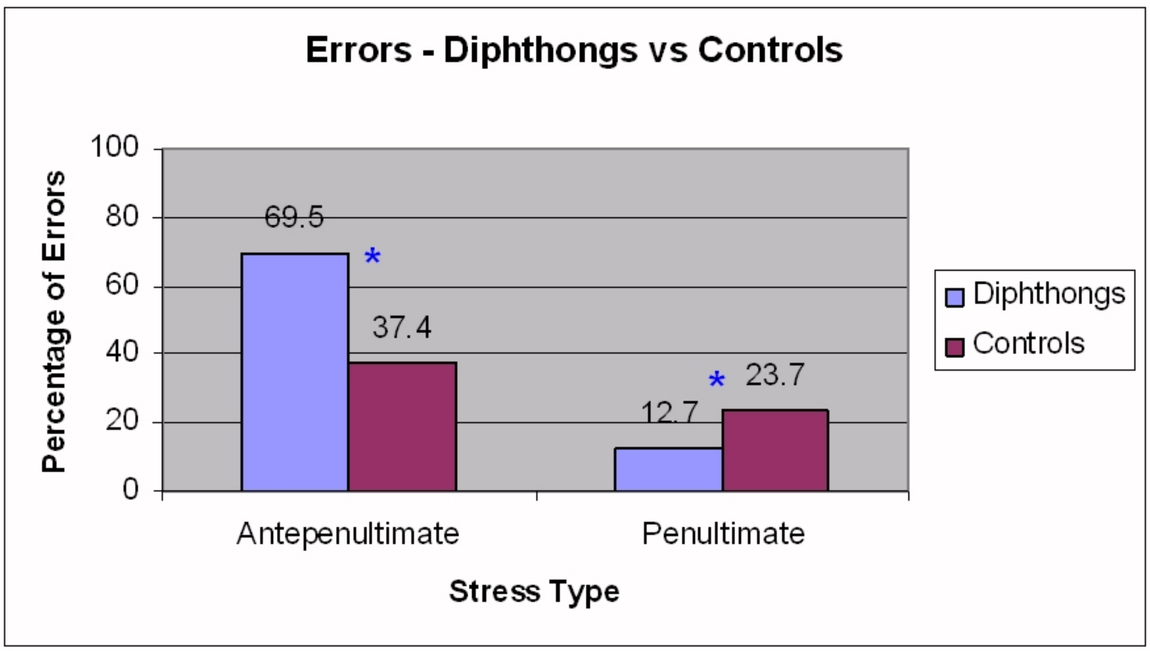

Figure **25**: Experiment 3: Error Rates – Diphthongs vs. Controls

In figure 25, we see a replication of the results from Experiment 2. The diphthong stimuli provoke significantly more errors in the antepenultimate condition than their control counterparts ($p<.001$). The finding is even more robust than in the previous experiment in that almost 70% of all proparoxytones with falling diphthongs in the penultimate syllable elicited production errors. This is an extremely high error rate which is indicative of the phonological constraint prohibiting stress placement to the left of a syllable containing a diphthong.

In the penultimate condition the opposite is the case. As expected, diphthong stimuli presented with paroxytone stress produce few errors when compared to monophthong controls ($p<.02$). Assuming that diphthongs create heavy syllables, and that heavy syllables attract stress, one would predict fewer errors for paroxytones with diphthongs in the penultimate syllable.

Turning our attention to the question of the discrepant behavior of the rising vs. falling diphthongs, two more 2 (diphthong type) x 2 (stress type) ANOVAs were run to examine how the two diphthong types regulate stress assignment in Spanish. The analysis of variance by participants and items for error rates revealed a significant main effect of diphthong type [$F1(1, 35) = 24.16$, $MSE = 0.22$, $p<.001$; $F2(1, 50) = 18.37$, $MSE = 0.15$, $p<.001$], a significant main effect of stress type [$F1(1, 35) = 156.34$, $MSE = 11.66$, $p<.001$; $F2(1, 50) = 933.13$, $MSE = 8.33$, $p<.001$], and a significant interaction of diphthong type and stress type [$F1(1, 35) = 10.21$, $MSE = 0.19$, $p<.003$; $F2(1, 50) = 15.64$, $MSE = 0.14$, $p<.001$]. These results thus also replicate the findings obtained in Experiment 2. Participants treat rising and falling diphthongs differently in both the antepenultimate and penultimate conditions. The following figure illustrates the results.

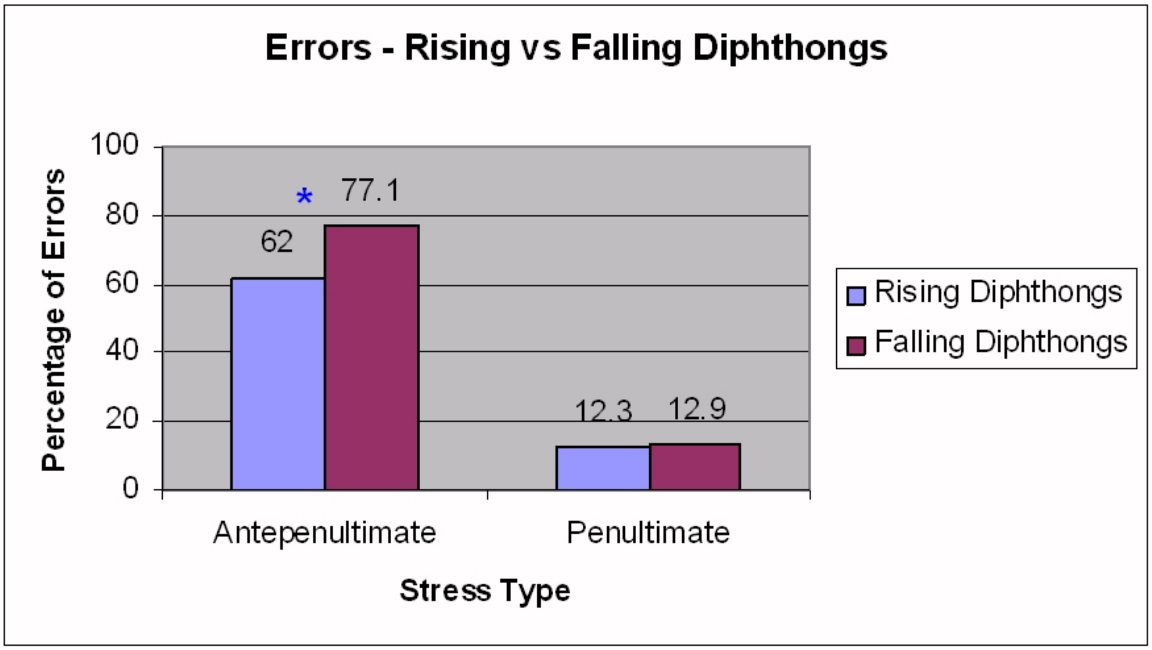

Figure **26**: Experiment 3: Error Rates – Rising Diphthongs vs. Falling Diphthongs

A significant difference was found between rising and falling diphthongs in the antepenultimate condition ($p<.001$). Again, these results replicate the findings from Experiment 2 in a very robust fashion. In general, the diphthong stimuli produce extremely high error rates in the antepenultimate condition. However, most importantly, the falling diphthongs provoke significantly more errors. In the penultimate condition, where both diphthong types are phonologically licit and predicted to attract stress, we find far lower error rates with no significant difference between the rising and falling diphthongs ($p>.99$).

These data also reflect the cognitive reality of the theoretical three-syllable window. Diphthongs create heavy syllables which attract stress, as illustrated by the lower error rate in the penultimate condition. Diphthongs also shrink the three-syllable window to one of two syllables. This is clearly displayed in the antepenultimate condition in which stress is assigned outside of the reduced two-syllable window. The latter condition exhibits extremely high error rates which are indicative of the proscribed nature of the antepenultimate condition.

Reaction time data, as in Experiment 2, are a weaker tool to examine the interaction of phonologically illicit syllable structures and stress assignment. As is suggested by the previous figures, the nonwords presented in Experiment 3 provoked so many errors that the amount of remaining accurate trials was greatly reduced. In the case of Experiment 3, five participants were eliminated from reaction time analyses comparing diphthongs and the controls, and seventeen were eliminated from the analyses comparing rising and falling diphthongs, due to empty cells. This, of course, significantly reduces the statistical power of the analyses. Nonetheless, the reaction times of the remaining participants were submitted to separate 2 x 2 repeated measures ANOVAs.

The first analyses by participants and items of reaction time data comparing diphthongs and the monophthong controls found a significant main effect of word type [$F1(1, 30) = 12.34$, $MSE = 155002$, $p<.002$; $F2(1, 105) = 12.98$, $MSE = 135748$, $p<.001$], a main effect of stress type that was significant by subjects [$F1(1, 30) = 7.58$, $MSE = 99341$, $p<.01$; $F2(1, 105) = 0.20$, $MSE = 2785$, $p>.65$], and an interaction of word type and stress type that was significant by participants[$F1(1, 30) = 4.68$, $MSE = 53161$, $p<.04$; $F2(1, 105) = 0.59$, $MSE = 8308$, $p>.44$]. Participants treated diphthongs and monophthongs differently in both the antepenultimate and penultimate condition. The following figure illustrates the reaction time data comparing diphthongs and their controls in each stress condition.

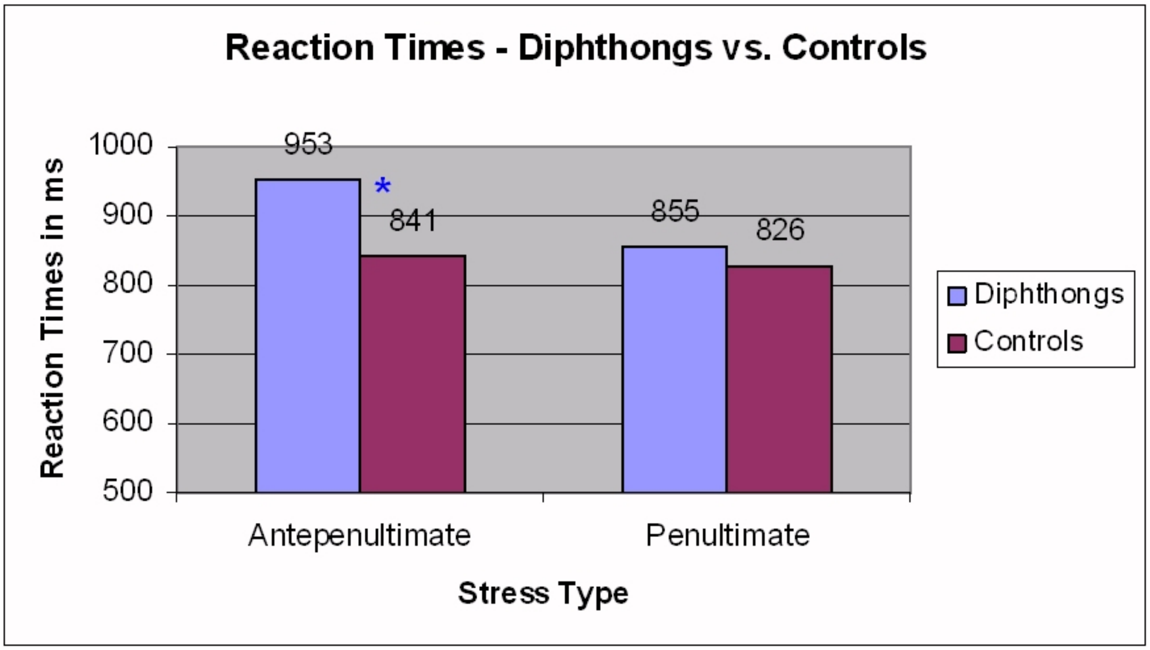

Figure **27**: Experiment 3: Reaction Times – Diphthongs vs. Controls

Despite having to eliminate five of the participants from the reaction time analyses due to empty cells caused by high error rates, there is still a clear effect of word type modulating stress assignment. In the antepenultimate condition, we see that participants produced diphthongs with significantly longer naming latencies ($p<.002$). In the penultimate condition, while diphthongs stimuli are pronounced with longer latencies on average, there is no significant difference between diphthong and monophthong items ($p>.70$).

These data in conjunction with the results from Experiment 2 indicate that the three syllable window is a definite cognitive reality in the mind of modern speakers of Spanish. Both accuracy and latency analyses of controlled diphthong and monophthong nonwords with participants from two very distinct dialects of Spanish evidence the same pattern of results. There is a strong interaction of word type and stress type which indicates that the synchronic phonology of Spanish prohibits proparoxytones with diphthongs in the penultimate syllable while permitting monophthongs in the penult. This restriction does not appear when diphthongs or monophthongs are presented with paroxytone stress. These data do not contradict approaches to Spanish stress theory that argue in favor of the continued importance of weight sensitivity in the modern language.

However, as discussed earlier, the difference in behavior between the rising and falling diphthongs is a much more finely-grained effect. In Experiment 2 the finding emerged robustly in the error rate data. In this experiment, the error rates are even higher than in the previous experiment, and again the robust effect of diphthong type found emerges in the error rates. As in Experiment 2, the effect is not found in the reaction time analyses. 2 (diphthong type) x 2 (stress type) repeated measures ANOVAs by participants and items found no significant main effect of diphthong type [$F1(1, 18) = 1.45$, $MSE = 5856$, $p > .24$; $F2(1, 50) = 0.82$, $MSE = 10550$, $p > .37$], a main effect of stress type that was significant by subjects [$F1(1, 18) = 12.38$, $MSE = 145890$, $p < .003$; $F2(1, 50) = 0.48$, $MSE = 10074$, $p < .49$], and no significant interaction of diphthong type and stress type [$F1(1, 18) = 1.39$, $MSE = 5240$, $p > .25$; $F2(1, 50) = 0.55$, $MSE = 11713$, $p > .46$]. Participants reacted differently to proparoxytones and paroxytones but made no distinction between diphthong types. The following figure illustrates the reaction time results for rising and falling diphthongs by stress condition.

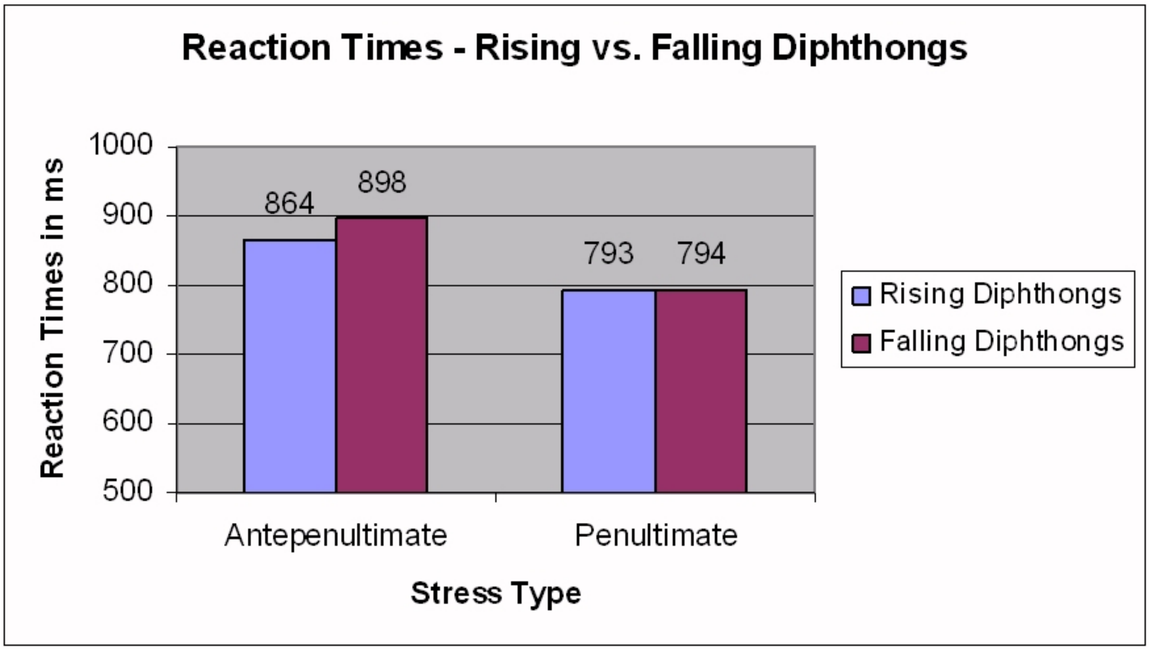

Figure **28**: Experiment 3: Reaction Times – Rising Diphthongs vs. Falling Diphthongs

The general impression given by figure 28 is the pattern we expect to find. In the antepenultimate condition, participants appear to pronounce falling diphthongs more slowly. In the penultimate condition the difference between latencies for rising and falling diphthongs is very slight. These differences are not significant (antepenultimate condition=$p>.34$, penultimate condition=$p>.99$). While the difference between diphthong and monophthong stimuli produces an effect that is sufficiently vigorous to survive the elimination of seventeen participants and the loss of large amounts of latency data due to high error rates, the difference between diphthong types is too subtle. The distinct treatment of the rising and falling diphthongs is straightforward in the accuracy data, and it is descriptively present in the reaction time data. However, the distinction appears to be too delicate to maintain statistical force under the current paradigm.

Finally, filler data were submitted to 2 (frequency) x 3 (stress type) repeated measures ANOVAs to test for frequency effects among the real words presented in the experiment. Results find a significant main effect of frequency [$F1(1, 35) = 70.04$, $MSE = 228998$, $p<.001$; $F2(1, 134) = 57.23$, $MSE = 152566$, $p<.001$], a main effect of stress type that was significant by participants [$F1(2, 70) = 3.82$, $MSE = 5317$, $p<.03$; $F2(2, 134) = 1.68$, $MSE = 4469$, $p>.19$], and no significant interaction of frequency and stress type [$F1(2, 70) = 0.93$, $MSE = 852$, $p>.39$; $F2(2, 134) = 0.45$, $MSE = 1195$, $p>.63$]. These results evidence a clear frequency effect, as illustrated in the following figure.

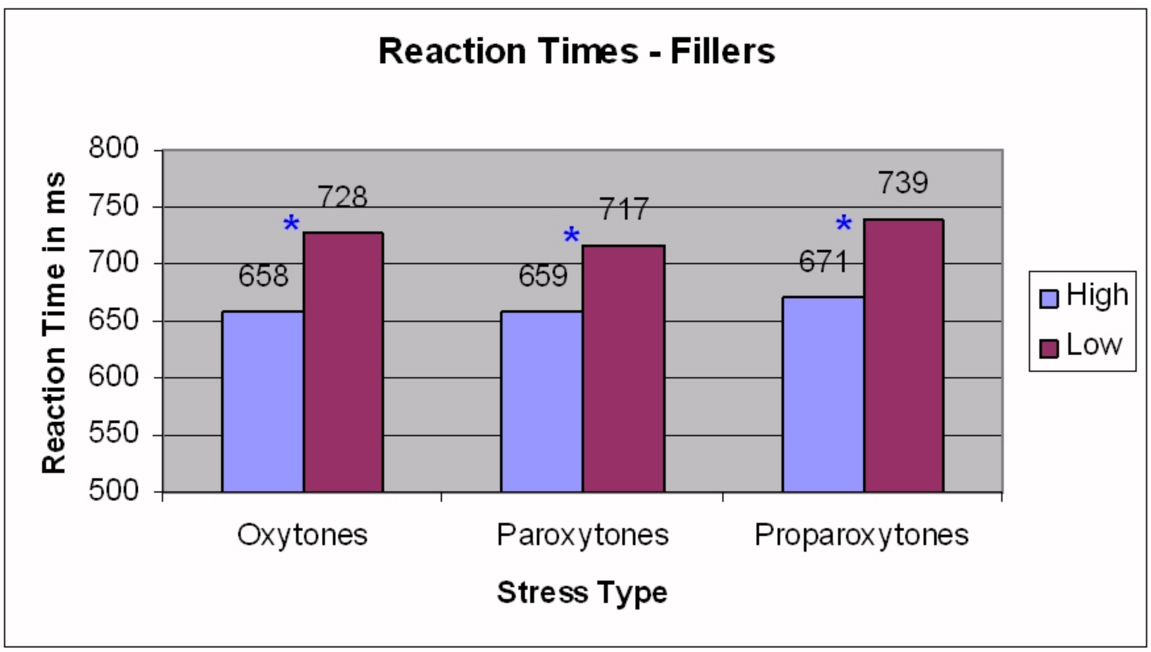

Figure **29**: Experiment 3: Reaction Times - Fillers

Here, the main effect of stress type is driven by the difference between proparoxytones and paroxytones of low frequency ($p<.04$). There is no significant difference between high frequency items by stress condition. Low frequency real words were produced significantly more slowly than high frequency words in each stress condition ($p<.001$). Therefore, we can preclude that the participants attended to the naming task given that the well established frequency effect among the real words was obtained.

Experiment 3 shows that the findings obtained in Experiment 2 are not a reflection of one particular speech community but rather a manifestation of shared knowledge among native speakers of Spanish. Despite numerous differences between the *jiennense* and *poblano* dialects of Spanish (cf. Zamora Vicente, 1960; Penny, 2000; Canfield, 1981; Lope Blanch, 1990; Lipski, 1994), the intuitions speakers of Spanish share about the possible and impossible word shapes in their native language appear to be a deep level of understanding below the surface-level variation. The differences found between the rising and falling diphthongs add to our understanding of the finely-grained ways in which we are able to track consistencies and inconsistencies across the lexicon. The methodology adapted here from cognitive psychology provides an effective tool to probe phonological knowledge. Of particular importance is that the collection of behavioral data allows us to eliminate the effects of metalinguistic interpretation and introspection on the part of the speaker by tapping into a level of knowledge beyond what native speakers consciously know about their language. This allows for the emergence of potentially subtle yet real differences that are not accessible to off-line, introspective tasks.

Given the findings in the first three experiments, it is clear that native speakers cognitively represent gradience in their sensitivity to the way in which putative syllable weight modulates the phonotactics of stress placement. Experiment 1 established that native speakers react differently to patterns that are hypothesized to be synchronic prohibitions versus diachronic gaps. Experiments 1-3 have found and replicated findings that there is a continuum for degrees of weight in which falling diphthongs form the heaviest vowel quality, followed by rising diphthongs as slightly less heavy, and monophthongs as the lightest syllable nucleus.

In Chapter 5, I build on the findings of Experiments 2-3, with the goal of probing for the locus of the phonological encoding of the divergent behavior of the diphthongs in modulating stress placement. Specifically, Experiment 4 will manipulate the time course of lexical processing as a means of locating the effects of syllable structure and stress encoding in psycholinguistic models of speech production. This will afford an integrated, cross-disciplinary approach to enhance our understanding of both how phonological constraints are represented and how they are processed by native speakers of Spanish.

Chapter 5

Experiment 4: Delayed Naming Task

5.1 Introduction

The purpose of Experiment 4 is to probe the time course of the diphthong effect found in Experiment 2 and replicated in Experiment 3. The findings of the previous two experiments indicate that native speakers of Spanish maintain distinct cognitive representations of the phonotactics of rising and falling diphthongs and their modulation of stress placement in the language. In a more general sense, the data show that syllable structure constrains stress placement in Spanish. What is not clear, however, is where this information is encoded in the time course of lexical processing.

Psycholinguistic models of speech production differ as to their account of syllable-level encoding. As discussed in Chapter 1, Dell's model posits that syllabic information is stored with the lexical item at the word level. The opposing view, adopted by Levelt and colleagues, is that syllables are constructed online at the form level of lexical processing. Experiment 4 employs the delayed naming task in order to investigate the time course of syllable encoding and its interaction with stress assignment in Spanish.

Given that the delayed naming task arguably excludes lexical effects, any remaining effects evidenced by the task are attributed to the phonological/phonetic encoding of the word form (Forster & Chambers, 1973; Monsell et al., 1989; Andrews & Heathcote, 2001). Therefore, a complete loss of the diphthong effect in Experiment 4 would suggest that syllable structure and stress encoding are stored with words in the mental lexicon. However, if the effect remains, we may attribute this effect to articulation processes. Also, any differences that arise between the treatment of rising or falling diphthongs at different delay intervals will be able to speak to the relative ease of encoding different phonotactic structures.

5.2 Participants

57 functionally monolingual adult speakers of Spanish participated in Experiment 4. All were undergraduate students at the University of Jaén, Spain. Two participants' data were excluded due to technical errors with the digital recordings rendering 55 total participants for analysis.

5.3 Materials

The materials for this experiment are the same as those described in section 3.3.

5.4 Procedure

Experiment 4 differs from the previous three experiments in that the participants performed a delayed naming task. As in the previous experiment, participants were tested in a quiet room. They were seated in front of a computer screen and one microphone. The microphone was connected to a digital recorder (Marantz PMD660 sampling at 48 kHz) that was directly connected to a PST Serial Response Box. The button box was connected to the computer and served as a voice key to cue the computer to record reaction times and to present the next stimulus when the participant spoke. The digital recorder was used to record the experimental session for subsequent coding of participant response accuracy.

In contrast with the previous experiments, the participants in Experiment 4 were instructed that they would be reading words one by one on a computer screen, and that they should read the words aloud as quickly and accurately as possible *when cued by a tone to speak*. Thus, participants withheld their pronunciation until they heard a tone. They were also informed that the tones may appear immediately with the word, or there may be a delay. The experimenter emphasized the need for the participant to prepare the word for pronunciation as soon as the word appeared so that they could pronounce the word immediately when the tone sounded, but not before. Following the instructions, the participants performed a practice session consisting of 12 items which allowed them to become accustomed to the task.

Each trial began with a fixation point (+) which was presented in the middle of the screen in black font on a white background. This was done to focus the participants' attention to the center of the screen and to cue them that the word they needed to pronounce was about to appear. The fixation point remained on the screen for 500 milliseconds and was followed a blank screen for 500 milliseconds. Following these 1000 milliseconds, the stimulus was presented. At delays of either 0, 600 or 1200 ms a short tone was played to cue the participants to pronounce the word. These intervals were chosen because data from the previous standard naming tasks showed that participants' reaction times to these stimuli were typically within the 600-700 ms range. Therefore, this delayed naming task consisted of three delay conditions: a 0 ms delay which is comparable to the standard naming task in which participants were instructed to pronounce the word immediately upon stimulus presentation, a 600 ms delay which represents the typical time it takes participants to produce the forms, and a longer delay of 1200 ms which requires participants to wait twice the typical amount of time before pronouncing the word. These differences allow the participants more or less time to encode phonological/phonetic information before cueing articulation (Laganaro & Alario, 2006). Upon hearing the tone, participants pronounced the word, and their speech triggered the voice key. This caused the stimulus on the screen to disappear and be replaced with a blank white screen. If the participant committed an error, the experimenter cued the computer to present an error screen. The word "Error" appeared in red letters in the upper left-hand corner of the screen. The purpose of this error message was to alert the participant to the fact that they had committed an error and to focus them on the task. After the error message, another blank screen appeared for 1000 milliseconds until the following trial began.

Once the participants completed the practice trials, the experimental stimuli were presented in random order with random delays over three experimental blocks with a 30 second break between them. The following diagram illustrates the delayed naming task as it was used in this experiment.

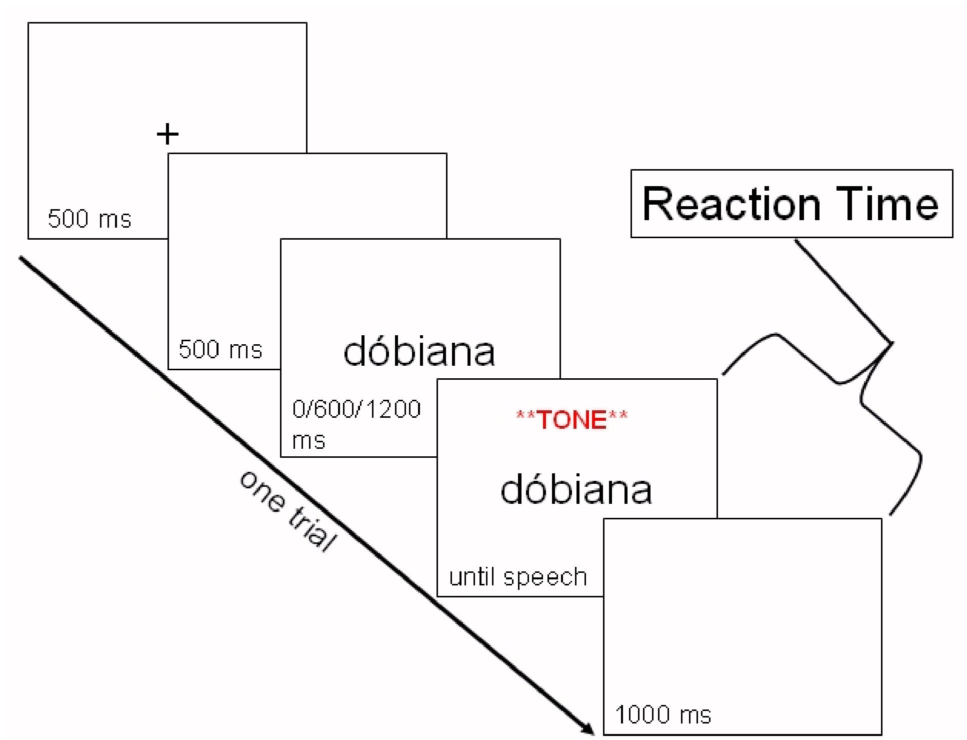

Figure **30**: Procedure for Experiment 4 (delayed naming task)

5.5 Data Analyses

Both error rate analyses and reaction time analyses were conducted on the data from this experiment. Native or near-native listeners (including the author) coded the errors manually using digital recordings of the experimental sessions. Errors comprised three categories: production errors, tone errors, and technical errors. Production errors consisted of instances in which participants misspoke the stimulus. Examples of this type of error were stress shifts to another syllable, changes in the segmental makeup of the word, or disfluencies in which words were not finished or contained pauses. Tone errors were instances in which the participants read the stimulus presented on the screen before the tone prompted articulation. Technical errors were those in which the microphone or the response box failed. These types of errors resulted in either extremely long reaction times, such as when the microphone did not trigger the voice key after articulation, or extremely fast reaction times, such as when the voice key triggered the next stimulus prematurely before the participant actually spoke.

The data were cleaned as follows. First, technical and tone errors were eliminated from all analyses. Production errors were included in error rate analyses, and excluded from reaction time analyses. All reaction time latencies were cleaned to exclude any responses faster than 200 ms and slower than 2000 ms. Subsequently, any latencies 2.5 standard deviations above or below each individual's mean reaction time were also eliminated.

Both reaction time and error data were submitted to separate repeated measures ANOVAs by participants ($F1$) to test for variance across the population and by items ($F2$) to test for variance across the stimuli set. Main effects and interactions were followed up by Tukey's HSD post-hoc tests to identify any independent effects across the experimental categories and conditions. Results are discussed below.

5.6 Results and Discussion

ANOVAs by participants and items of error rates comparing diphthongs to their monophthong controls found a significant main effect of stress type (proparoxytone vs. paroxytone) [$F1(1, 54) = 34.64$, $MSE = 2.08$, $p<.001$; $F2(1, 110) = 91.05$, $MSE = 2.08$, $p<.001$], a significant main effect of word type (diphthong vs. monophthong) [$F1(1, 54) = 67.84$, $MSE = 2.13$, $p<.001$; $F2(1, 110) = 61.96$, $MSE = 2.00$, $p<.001$], a significant main effect of delay (0 ms, 600 ms, 1200 ms) [$F1(2, 108) = 42.68$, $MSE = 0.78$, $p<.001$; $F2(2, 220) = 32.83$, $MSE = 0.66$, $p<.001$], a significant interaction of delay and word type [$F1(2, 108) = 4.08$, $MSE = 0.05$, $p<.02$; $F2(2, 220) = 5.44$, $MSE = 0.11$, $p<.01$], an interaction of delay and stress type that was significant by items [$F1(2, 108) = 1.20$, $MSE = 0.02$, $p>.30$; $F2(2, 220) = 3.16$, $MSE = 0.05$, $p<.05$], a significant interaction of word type and stress type [$F1(1, 54) = 90.24$, $MSE = 4.57$, $p<.001$; $F2(2, 220) = 187.88$, $MSE = 4.29$, $p<.001$], and an interaction of delay, word type, and stress type that was significant by participants and approaches significance by items [$F1(2, 108) = 3.98$, $MSE = 0.05$, $p=.02$; $F2(2, 220) = 2.88$, $MSE = 0.05$, $p=.058$]. The most important finding from this analysis is the significant interaction of delay, word type, and stress type. It is apparent that participants treated diphthongs and monophthongs differently when presented with proparoxytone and paroxytone stress. As in Experiments 1-3, this interaction is critical, because it shows that there is a difference in treatment of the critical items between the theoretically permitted and prohibited stress conditions. The reaction to the two word types also varied by delay. This interaction is key to the goal of tracking the time course of the phonological encoding. The following figure illustrates the error rates for the diphthong and monophthong control stimuli in the 0 ms delay condition.

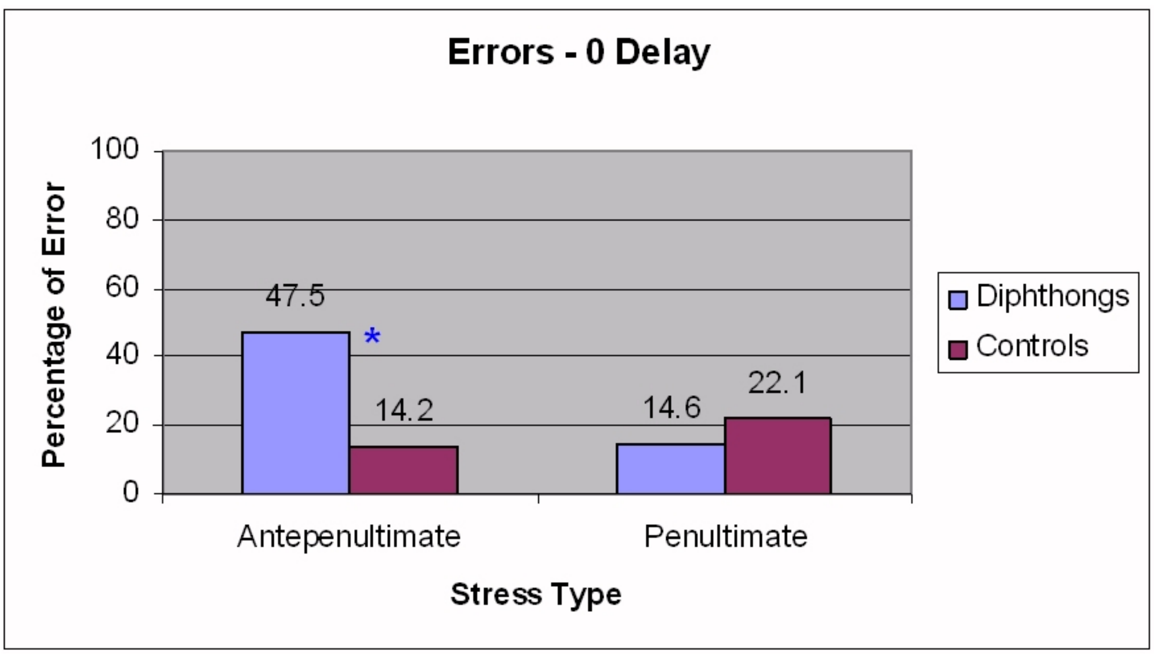

Figure **31**: Experiment 4: Error Rates – Diphthongs vs. Controls – 0 ms Delay

In a sense, the 0 ms delay condition should replicate the findings of the previous experiments in that the standard naming task consists of naming with no delay. While the overall task demands of a delayed naming experiment may influence participant reactions within the condition with no delay, we see clearly in figure 31 that Experiment 4 also evidences a clear prohibition on proparoxytones with diphthongs in the penultimate syllable. The diphthong stimuli provoked significantly more errors than monophthong controls in the antepenultimate condition (p<.001). In the penultimate condition, the difference between word types approaches significance (p>.07). In previous chapters, the difference between the diphthong stimuli and their monophthong controls proved significant in the penultimate condition. Monophthong controls evidenced a higher rate of error than the diphthong stimuli in Experiments 2 and 3. We attributed this significant difference to the fact that the monophthong controls are licit forms in either the antepenultimate or penultimate condition, and therefore they were more prone to variance in responses. The diphthong stimuli, in contrast, are predicted to attract stress in the penultimate condition and, therefore, provoke fewer errors. The higher rate of error for the control stimuli is also likely to be an artifact of the unusually high number of proparoxytone forms throughout the experiment. Nonetheless, the lack of a significant difference in the penultimate condition in Experiment 4 is not problematic for our predictions. While weight-sensitive accounts of Spanish stress are able to explain the significant differences found in Experiment 2 and 3, no significant difference between the diphthong and monophthong stimuli is not surprising, given they are both phonologically licit phonotactic sequences when presented with paroxytone stress. The results from the antepenultimate condition, however, support the finding of the previous experiments. The theoretically proscribed forms evidence higher rates of error. This offers added support for a quantity-sensitive account of Spanish stress.

This same pattern is evidenced across all three delays. The following figures illustrate the error rates for the 600 ms and 1200 ms delays.

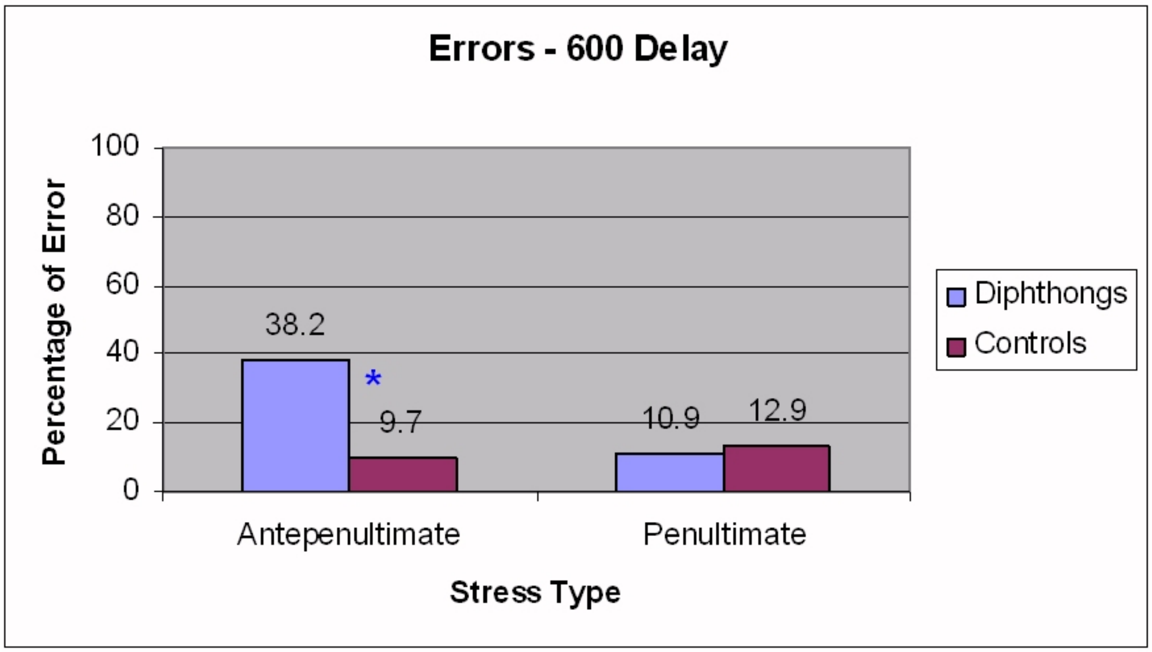

Figure **302**: Experiment 4: Error Rates – Diphthongs vs. Controls – 600 ms Delay

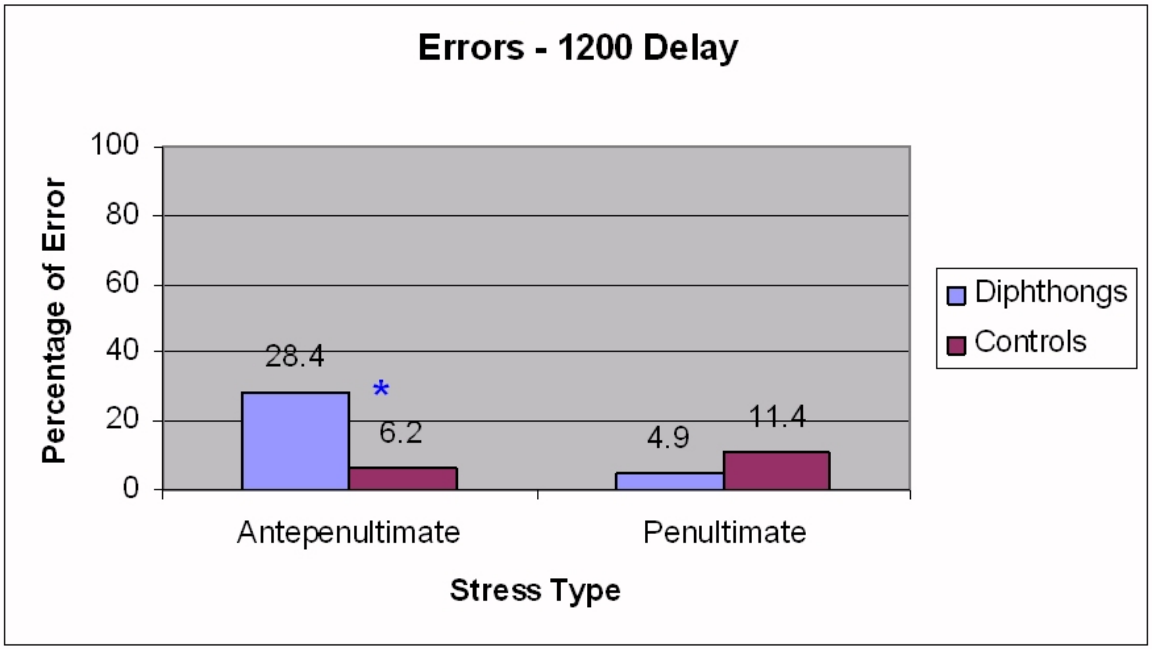

Figure **33**: Experiment 4: Error Rates – Diphthongs vs. Controls – 1200 ms Delay

As would be expected with more preparation time to encode the phonological form of the nonwords, the overall error rates decrease across all conditions as the delay time increases. Nevertheless, the difference between diphthong and control error rates in the antepenultimate condition is very robust across all delays ($p<.001$). The difference between the two word types in the penultimate condition is also not significant in the 600 ms and 1200 ms delay conditions (600 = $p>.99$, 1200 = $p>.18$). Figures 31-33 clearly replicate the results from the previous two experiments. They are indicative of a strong constraint on the Spanish stress system in which diphthongs in the penultimate syllable prohibit proparoxytone stress. The overall decrease of error rates suggests that the participants performed the delayed naming task appropriately. More preparation time allows speakers to prepare the words before speaking. As the amount of preparation time increases, so does the accuracy of their production.

In order to test for the differential patterning of rising and falling diphthongs evidenced in Experiments 2 and 3, error rates were submitted to separate 2 (stress type) x 2 (diphthong type) x 3 (delay) repeated measures ANOVAs. Results show a significant main effect of delay [$F1(2, 108) = 28.19$, $MSE = 0.98$, $p<.001$; $F2(2, 108) = 34.27$, $MSE = 0.65$, $p<.001$], a significant main effect of stress type [$F1(1, 54) = 73.57$, $MSE = 12.99$, $p<.001$; $F2(1, 54) = 496.28$, $MSE = 6.63$, $p<.001$], a significant main effect of diphthong type [$F1(1, 54) = 15.15$, $MSE = 0.78$, $p<.001$; $F2(1, 54) = 22.72$, $MSE = 0.54$, $p<.001$], a significant interaction of delay and stress type [$F1(2, 108) = 4.23$, $MSE = 0.11$, $p<.02$; $F2(2, 108) = 7.65$, $MSE = 0.13$, $p<.001$], no significant interaction of delay and diphthong type [$F1(2, 108) = 2.53$, $MSE = 0.07$, $p>.08$; $F2(2, 208) = 2.67$, $MSE = 0.05$, $p>.07$], an interaction of stress type and diphthong type that approaches significance by participants and is significant by items [$F1(1, 54) = 3.42$, $MSE = 0.17$, $p>.06$; $F2(1, 54) = 16.86$, $MSE = 0.23$, $p<.001$], and no significant interaction of delay, stress type, and diphthong type [$F1(2, 108) = 0.89$, $MSE = 0.03$, $p>.41$; $F2(2, 108) = 2.39$, $MSE = 0.04$, $p>.09$]. The main results of these ANOVAs do not evidence differential treatment of rising and falling diphthongs across stress conditions, although the effect approaches significance. The crucial interaction of all three independent variables (delay, stress type, and diphthong type) also fails to prove significant. Therefore, initially, it appears that Experiment 4, with its additional delay variable, fails to replicate the diphthong effect evidenced in Experiments 2 and 3. However, pairwise comparisons from post-hoc analyses do suggest differential treatment of the diphthong types in the 600 ms delay condition, and a distinct pattern for rising and falling diphthongs across the three delay conditions. The following figure illustrates the error rates for rising vs. falling diphthongs in the 0 ms delay.

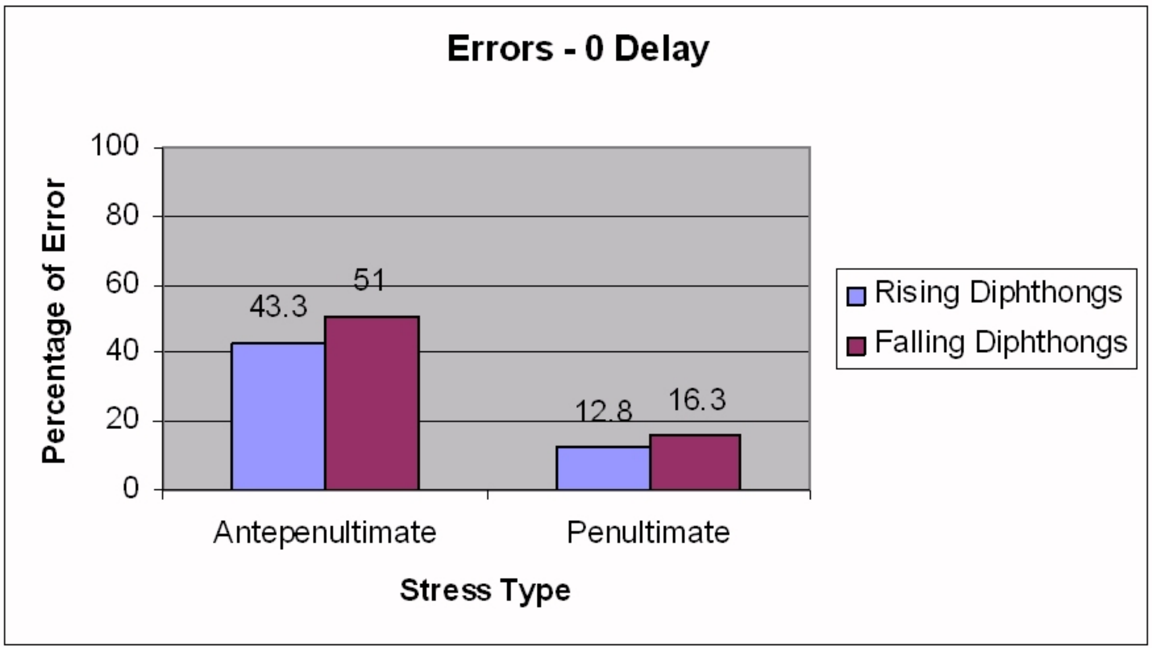

Figure **34**: Experiment 4: Error Rates – Rising Diphthongs vs. Falling Diphthongs – 0 ms Delay

Although descriptive statistics show that the falling diphthongs provoke a higher rate of errors in the antepenultimate condition, this difference is not statistically significant ($p>.57$). This result is surprising given the robust difference between the two diphthong types confirmed in Experiments 2 and 3. Arguably, the 0 ms delay should be comparable to the standard naming task, since the standard naming task consists of the immediate articulation of the target item upon visual presentation. In the 0 ms delay condition, the tone that cued the participants to produce the word was presented at the same time as the visual presentation of the stimulus. However, in the standard naming task, participants know that they need to produce the word as soon as it appears on the screen. In the delayed naming task, participants anticipate a delay. If there is no delay, i.e. the tone sounds immediately, participants may arguably react differently because the lack of a delay is unexpected. This variability in the task likely added noise to the data in Experiment 4. While the diphthong effect has proven robust when diphthong stimuli are tested only with control items and filler stimuli as in Experiments 2 and 3, we also know that the effect is subtle. In Experiment 1 the effect did not surface when too few items were presented with various other phonotactic configurations. Therefore, the lack of a diphthong effect in 0 ms may arguably be an artifact of the task conditions.

Under the assumption that the surprise of a 0 ms delay in a delayed naming task slows down participant responses, separate ANOVAs were conducted on a subset of the faster participants. The reasoning was, if the faster participants in this experiment had faster reaction times on correct trials because they were less influenced by the lack of a delay, perhaps their behavior in the 0 ms delay would more closely resemble the results found in the standard naming task. 29 of the original 55 participants were included in the second analysis. They were chosen by performing a median split on the total participant group. Those participants with a mean reaction time higher than the group mean were included. The 2 (stress type) x 2 (diphthong type) x 3 (delay) repeated measures ANOVA found a significant main effect of stress type [$F1(1, 28) = 38.79$, MSE =

6.15, $p<.001$; $F2(1, 54) = 213.44$, $MSE = 5.82$, $p<.001$], a significant main effect of diphthong type [$F1(1, 28) = 8.62$, $MSE = 0.36$, $p<.01$; $F2(1, 54) = 14.21$, $MSE = 0.49$, $p<.001$], a significant main effect of delay [$F1(2, 56) = 22.53$, $MSE = 0.62$, $p<.001$; $F2(2, 108) = 17.48$, $MSE = 0.73$, $p<.001$], no significant interaction of delay and stress type [$F1(2, 56) = 0.57$, $MSE = 0.01$, $p>.56$; $F2(2, 108) = 0.30$, $MSE = 0.01$, $p>.73$], no significant interaction of delay and diphthong type [$F1(2, 56) = 1.63$, $MSE = 0.05$, $p>.20$; $F2(2, 108) = 1.43$, $MSE = 0.06$, $p>.24$], no significant interaction of stress type and diphthong type [$F1(1, 28) = 1.04$, $MSE = 0.05$, $p>.31$; $F2(1, 54) = 3.74$, $MSE = 0.10$, $p>.05$], and no significant interaction of delay, stress type, and diphthong type [$F1(2, 56) = 022$, $MSE = 0.01$, $p>.80$; $F2(2, 108) = 0.96$, $MSE = 0.04$, $p>.38$]. The results of the analyses of the faster participants replicate the findings of the general analyses in that there is no significant interaction of stress type and diphthong. While participants treated both diphthongs different, and both stress types differently, these two categories did not interact. The following figure illustrates the error rate data for the rising and falling diphthongs in the 0 ms delay.

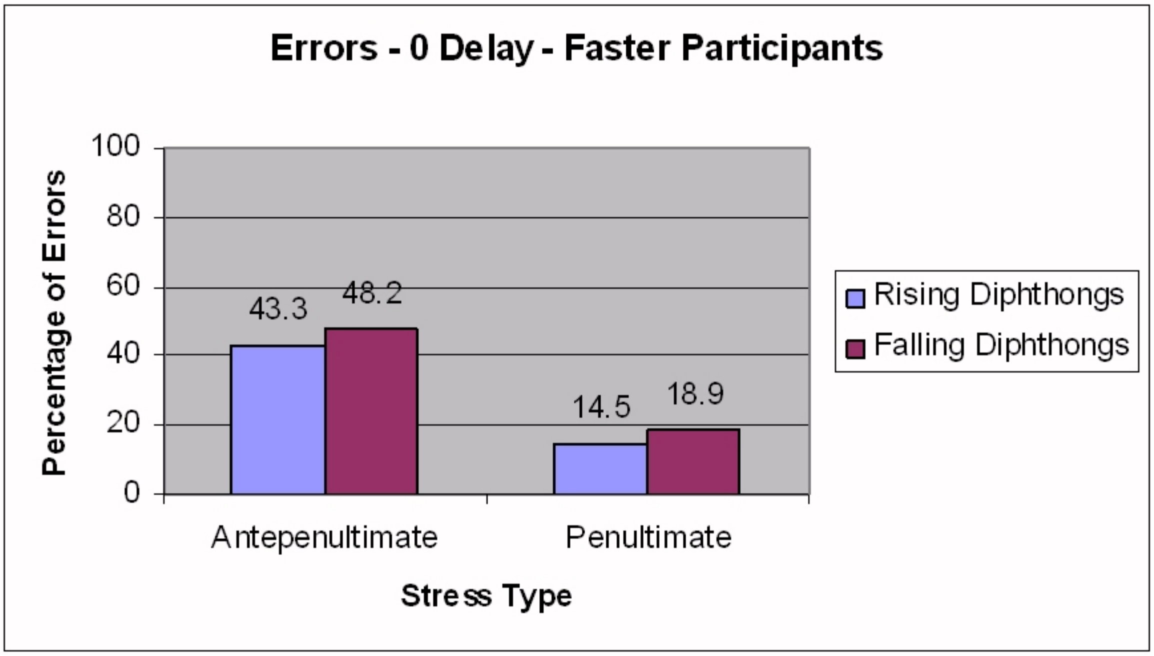

Figure 35: Experiment 4: Error Rates – 0 ms Delay – Faster Participants

The prediction that the faster participants would not be as affected by the 0 ms delay did not hold. Post-hoc tests show no significant difference between rising and falling diphthongs in the antepenultimate condition for the faster participants (*p*>.99). It appears that all participants were affected by the unexpected lack of a delay.

Nonetheless, the effect is not completely absent from the results. Experiment 4 does evidence the diphthong effect in the 600 ms delay condition. The following figure depicts the error rates for rising and falling diphthongs in the second delay.

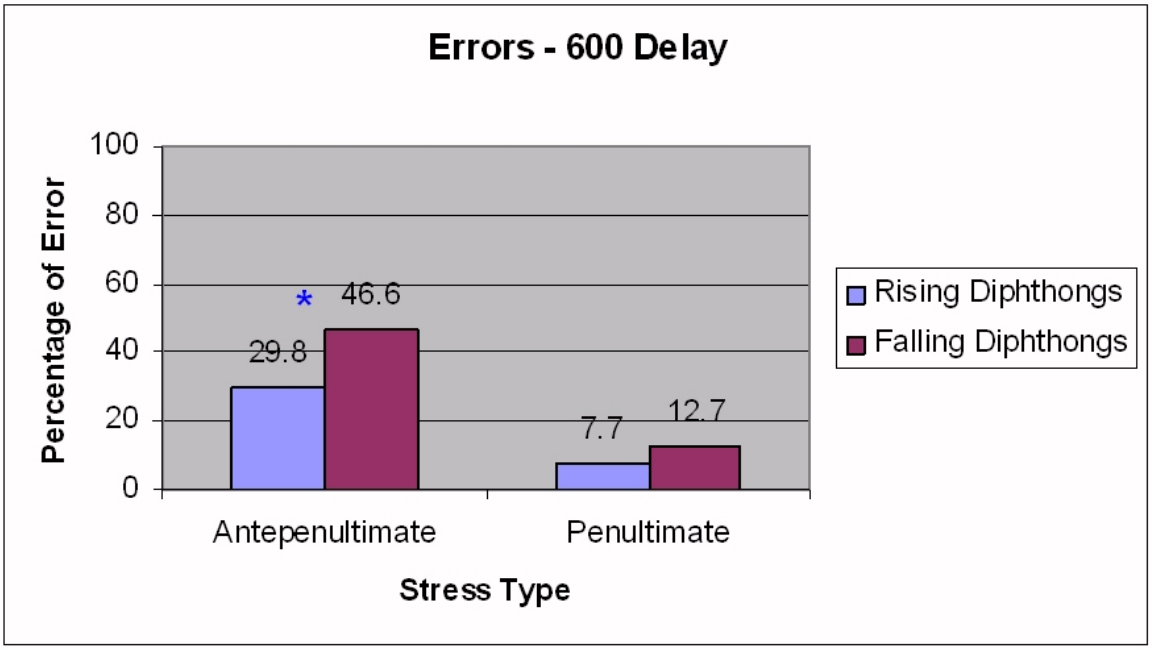

Figure **36**: Experiment 4: Error Rates – Rising Diphthongs vs. Falling Diphthongs – 600 ms Delay

In the antepenultimate condition of the 600 ms delay, participants committed more errors when presented with falling diphthong stimuli than rising diphthong stimuli (p<.001). The difference between the two diphthong types in the penultimate condition is not statistically significant (p>.95). Thus we see that the 600 ms delay does replicate the diphthong effect found in previous experiments.

Interestingly, however, the result disappears again in the 1200 ms delay, as illustrated in the figure below.

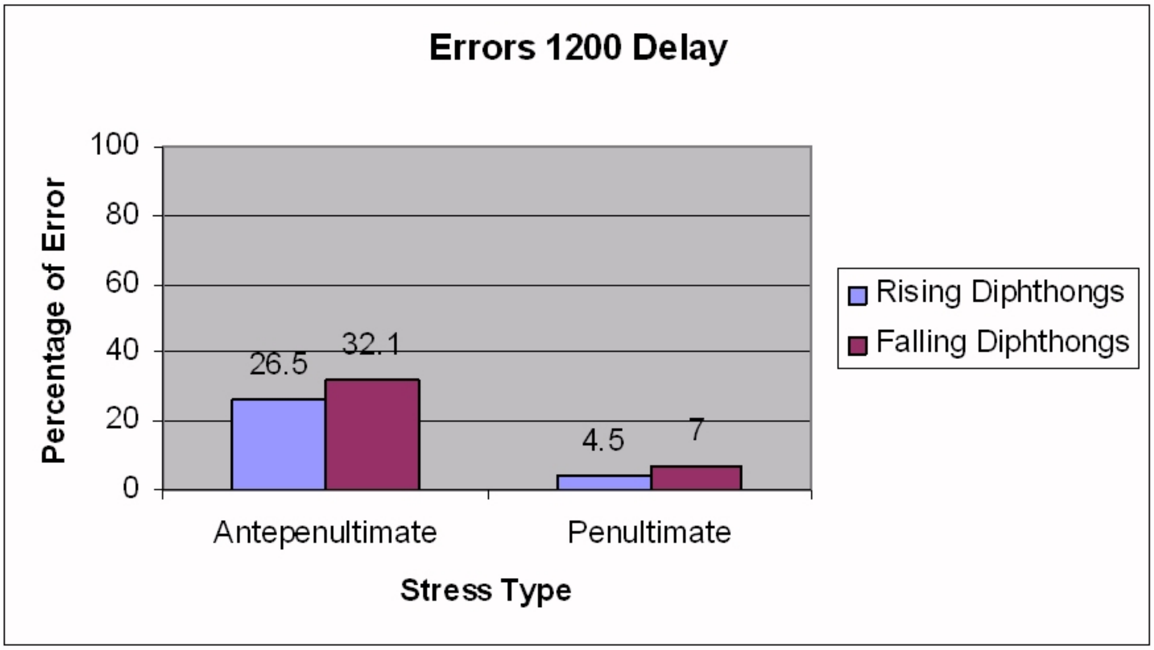

Figure 37: Experiment 4: Error Rates – Rising Diphthongs vs. Falling Diphthongs – 1200 ms Delay

Results reveal an overall decline in error rates, as is expected with the considerably long delay of 1200 ms. Post-hoc analyses of the 1200 ms delay condition exhibit no significant differences between diphthong types in either condition (antepenultimate = $p>.92$, penultimate = $p>.99$). In sum, as we have seen in the previous three figures and illustrated in the following graph, the diphthong effect does not appear in the 0 ms delay, appears in the 600 ms delay, and disappears again in the 1200 ms delay.

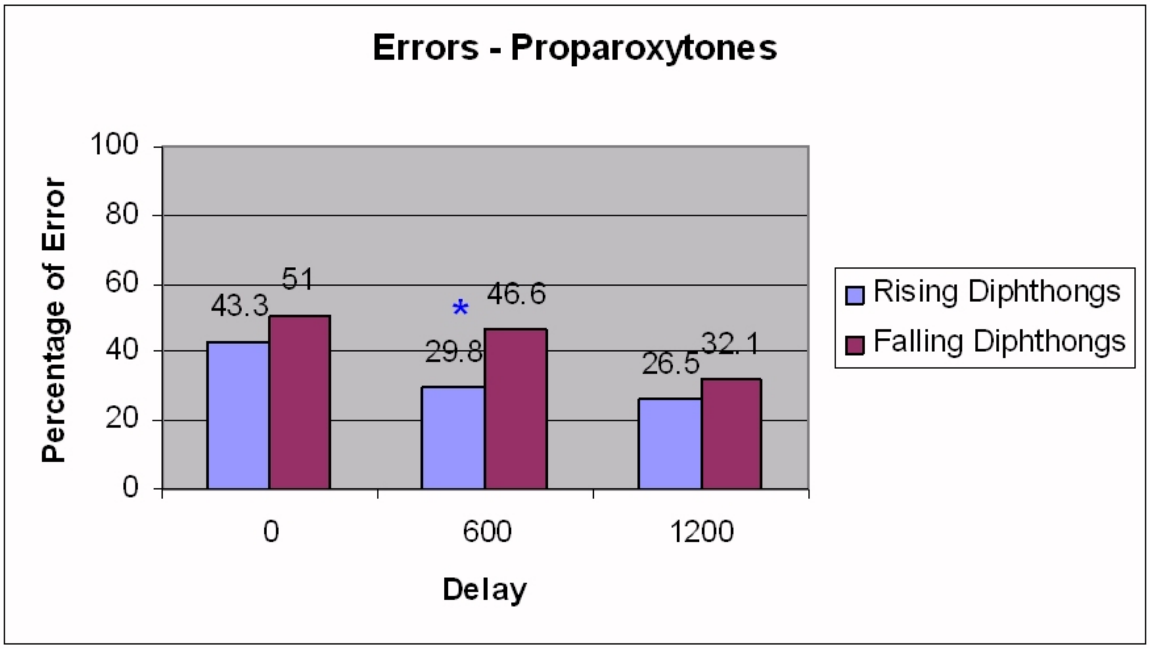

Figure 38: Experiment 4: Error Rates – Rising Diphthongs vs. Falling Diphthongs – Only proparoxytones across all delays

One possible account for this pattern of results is the following. The 0 ms delay, as discussed earlier, was a surprising condition for the participants. In order to speak, they expected to have to wait for a tone to appear *after* the word was presented. When the tone sounded *at the same time* as the word, this unexpected occurrence caused variability in the participants' reactions. Given the subtlety of the diphthong effect, this added noise in the task washed away the statistical effect from the data. Descriptively it is clear from figure 38 that falling diphthongs systematically provoked high error rates. However, the statistical significance of these differences varies. Therefore, when an actual delay did appear between the stimulus presentation and the tone, the participants reacted in the predicted manner. The diphthong effect surfaces in the 600 ms condition, as would be predicted theoretically and based on the findings from previous chapters of this dissertation. Yet the effect does not emerge in the 1200 ms delay. Arguably, 1200 ms is enough time for the participants to prepare to pronounce these difficult forms. Clearly, they do not perform perfectly; the participants still commit production errors 25.5-31.1% of the time. However, the preparation time allotted by a 1200 ms delay appears to be enough time for the fine-grained difference between the rising and falling diphthongs to disappear from the data.

A potentially clearer interpretation of the data becomes visible in the following figure.

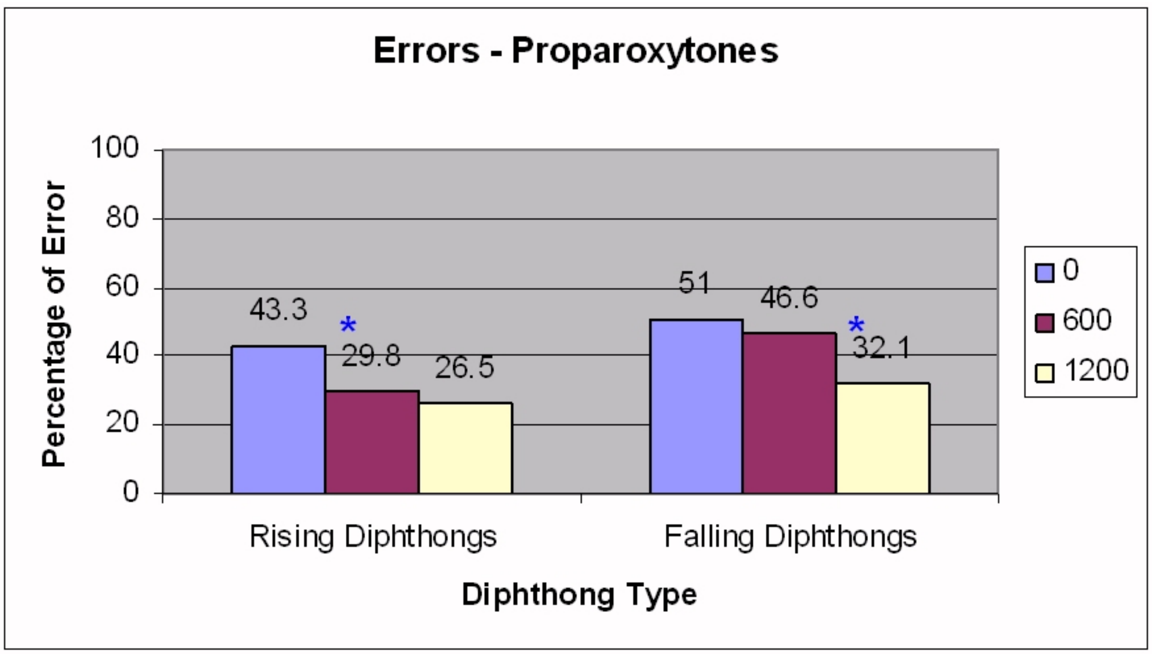

Figure **39**: Experiment 4: Error Rates – Rising Diphthongs vs. Falling Diphthongs – Time-course changes for each diphthong type

This figure tracks the changes over the delay conditions of each diphthong type. Participants reacted differently to the rising diphthongs in the 0 ms and 600 ms delays (p<.02). There is no significant difference between the error rates provoked by the rising diphthongs in the 600 ms and 1200 ms delays (*p*>.99). In contrast, the significant difference for the falling diphthongs appears between the 600 ms and 1200 ms delays (*p*<.01). There is no differential treatment of falling diphthongs between the 0 ms and 600 ms delays (*p*>.98). These data suggest that 600 ms is a sufficient amount of preparation time to facilitate the phonological/phonetic encoding of the rising diphthongs. However, the falling diphthongs require more preparation time for any facilitation to appear. The falling diphthongs do not evidence an effect of facilitation until the 1200 ms delay.

This patterning between the rising and falling diphthongs is consequent with what I have argued is the *lighter* nature of the rising diphthongs. That is, if we predict that the rising diphthongs should pattern as *less heavy*, then the strength with which they block the possibility of stress in the antepenultimate condition should be weaker than for the falling diphthongs. As discussed in Chapter 1, distributional evidence shows that the falling diphthongs pattern systematically as *heavy* syllables. Experiments 2 and 3 confirm this patterning experimentally. At first examination, it appeared that Experiment 4 did not replicate these results. However, figure 39 does illustrate that the falling diphthongs require twice as much preparation time as the rising diphthongs in order to evidence facilitation in preparation. These results accord with a weight sensitive interpretation of the Spanish stress system. Rising diphthongs are represented as *lighter* syllables, and therefore are less constrained in the antepenultimate condition. This weaker restriction is evidenced by a lower rate of error in Experiments 2 and 3. In Experiment 4, the time course of phonological preparation, as seen by the results across the delay conditions, suggests that the rising diphthongs are easier to encode given their earlier facilitation by shorter delays. This finding thus corroborates the findings of Experiments 2 and 3 within a different task, and provides support for the gradient representation of phonotactic constraints in the mind of native Spanish speakers.

Reaction time data were also submitted to separate repeated measures ANOVAs. First, 2 (stress type) x 2 (word type) x 3 (delay) ANOVAs were run to compare response latencies between diphthongs and their monophthong controls. Six participants were excluded from the analysis due to empty cells caused by the high error rates in the experiment. Results found a significant main effect of stress type [$F1(1, 48) = 11.31$, $MSE = 54223$, $p<.002$; $F2(1, 110) = 9.12$, $MSE = 29662$, $p<.01$], a significant main effect of word type [$F1(1, 48) = 46.51$, $MSE = 136101$, $p<.001$; $F2(1, 110) = 30.14$, $MSE = 126906$, $p<.001$], a significant main effect of delay [$F1(2, 96) = 323.71$, $MSE = 3120762$, $p<.001$; $F2(2, 220) = 904.05$, $MSE = 3672679$, $p<.001$], no significant interaction of delay and word type [$F1(2, 96) = 1.09$, $MSE = 3006$, $p>.33$; $F2(2, 220) = 0.38$, $MSE = 1550$, $p>.68$], no significant interaction of delay and stress type [$F1(2, 96) = 1.21$, $MSE = 3770$, $p>.30$; $F2(2, 220) = 2.79$, $MSE = 10110$, $p>.06$], a significant interaction of word type and stress type [$F1(1, 48) = 14.39$, $MSE = 34269$, $p<.001$; $F2(1, 110) = 6.67$, $MSE = 21671$, $p<.02$], and no significant interaction of delay, word type and stress type [$F1(2, 96) = 1.64$, $MSE = 4258$, $p>.19$; $F2(2, 220) =$

2.32, *MSE* = 8404, *p*>.10]. The results of these ANOVAs show that diphthongs and monophthongs were treated differently in both stress conditions. This difference does not vary statistically by delay. The following figure details the reaction time data for diphthongs and their controls in the antepenultimate condition across the three delays.

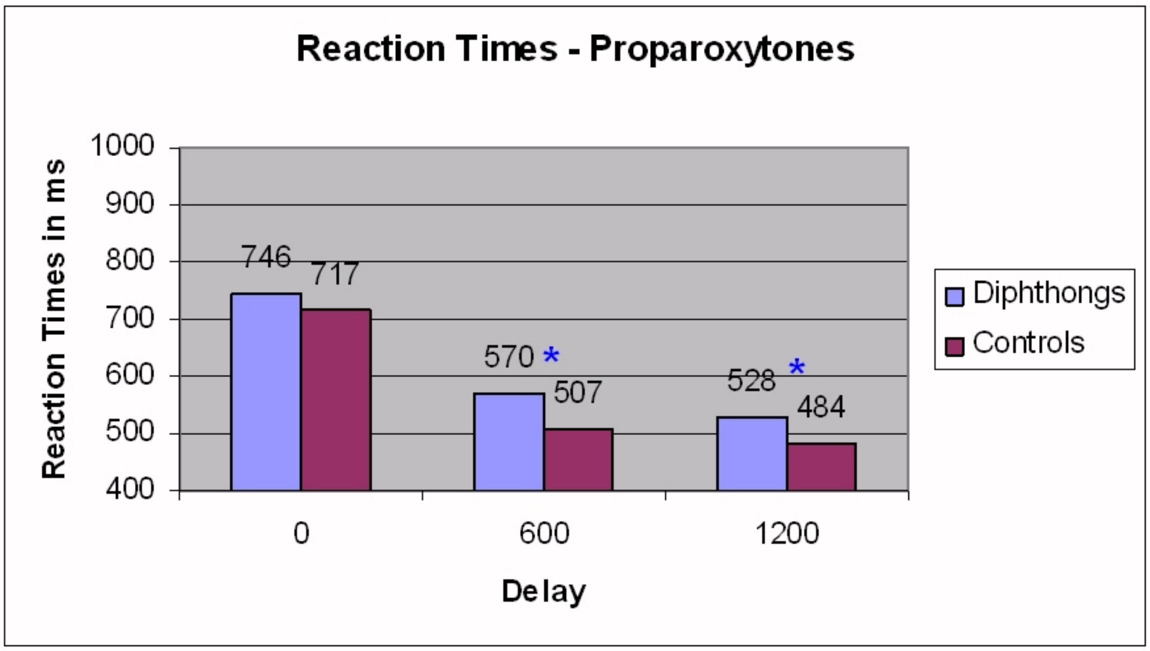

Figure **40**: Experiment 4: Reaction Times – Diphthongs vs. Controls – Only proparoxytones across all delays

Diphthongs evoked slower reaction times in the 600 ms (*p*<.001) and in the 1200 ms (*p*<.01) delays. While descriptively it appears that diphthongs were also pronounced more slowly in the 0 ms delay, the difference is not statistically significant (*p*>.19). The lack of an effect in the 0 ms delay supports the previous interpretation of the 0 ms delay for error rates. The difference between diphthongs in general and their monophthong controls disappears in the 0 ms condition but continues to be found in the 600 ms and 1200 ms delay conditions. This indicates that the 0 ms condition is not a viable condition for making comparisons here. The results from the 600 ms and 1200 ms delays, however, show that diphthongs were produced more slowly than the monophthong controls across the delay conditions. Since it is with proparoxytone stress that the diphthongs are predicted to be theoretically illicit forms, these results appear to support the findings from previous experiments as well. This is only true, of course, if there are no significant differences between the diphthongs and controls in the penultimate condition. The following figure depicts the diphthongs and controls across all three delays when they were presented with paroxytone stress.

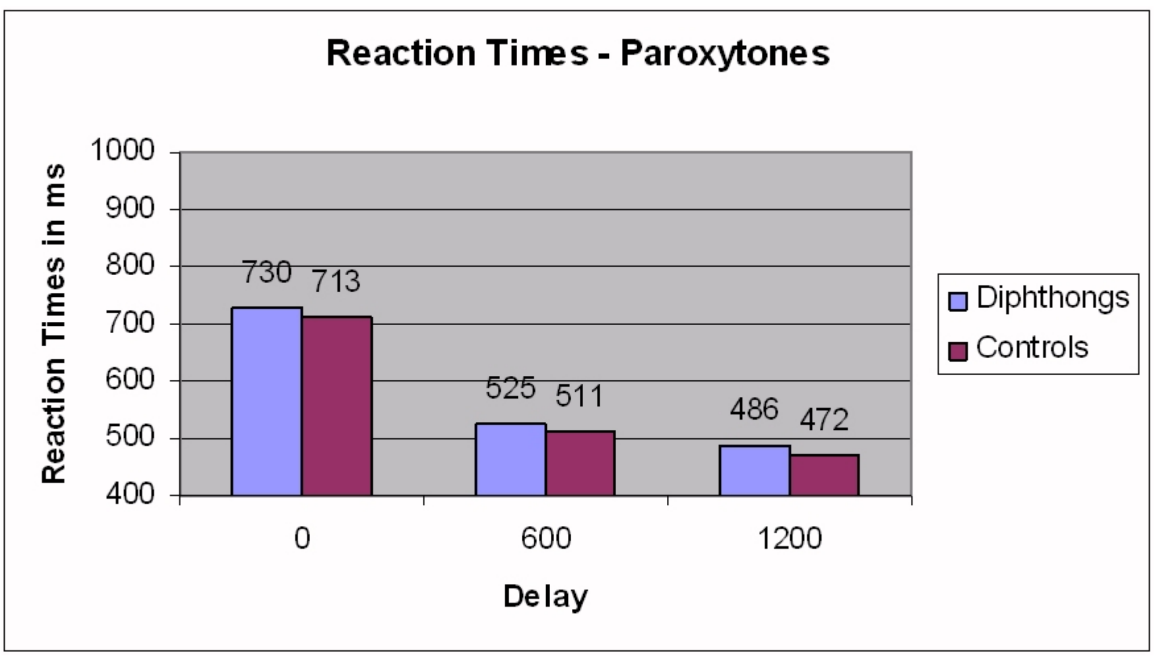

Figure **41**: Experiment 4: Reaction Times – Diphthongs vs. Controls – Only paroxytones across all delays

There were no significant differences between diphthongs and the control stimuli in the penultimate condition. This again accords with a weight-sensitive interpretation of the Spanish stress system.

Reaction time data comparing the two diphthong types were also submitted to separate 2 (stress type) x 2 (diphthong type) x 3 (delay) repeated measures ANOVAs. For this analysis 23 participants were excluded due to empty cells created by high error rates. Results find a significant main effect of stress type [$F1(1, 31) = 18.07$, $MSE = 147565$, $p<.001$; $F2(1, 54) = 18.04$, $MSE = 51021$, $p<.001$], no significant main effect of diphthong type [$F1(1, 31) = 0.78$, $MSE = 3469$, $p>.38$; $F2(1, 54) = 0.03$, $MSE = 113$, $p>.87$], a significant effect of delay [$F1(2, 62) = 110.21$, $MSE = 1980924$, $p<.001$; $F2(2, 108) = 468.86$, $MSE = 1808898$, $p<.001$], no significant interaction of delay and stress type [$F1(2, 62) = 0.55$, $MSE = 2795$, $p>.57$; $F2(2, 108) = 2.08$, $MSE = 8587$, $p>.12$], no significant interaction of delay and diphthong type [$F1(2, 62) = 0.86$, $MSE = 4834$, $p>.42$; $F2(2, 108) = 2.49$, $MSE = 9622$, $p>.08$], no significant interaction of stress type and diphthong [$F1(1, 31) = 0.93$, $MSE = 5994$, $p>.34$; $F2(1, 54) = 0.07$, $MSE = 197$, $p>.79$] and no significant interaction of delay, stress type and diphthong type [$F1(2, 62) = 0.57$, $MSE = 2264$, $p>.56$; $F2(2, 108) = 0.11$, $MSE = 453$, $p>.89$]. These results reveal no significant interactions of any kind. Given the high amount of error rates, over 41% of the participants' data had to be excluded from the statistical analyses. This diminished quantity was not sufficiently powerful to evidence the diphthong effect found in the accuracy data. In this sense, Experiment 4 fits the pattern of results observed in Experiments 2 and 3. The following figure illustrates the reaction times for rising and falling diphthongs in the antepenultimate condition across the three delays.

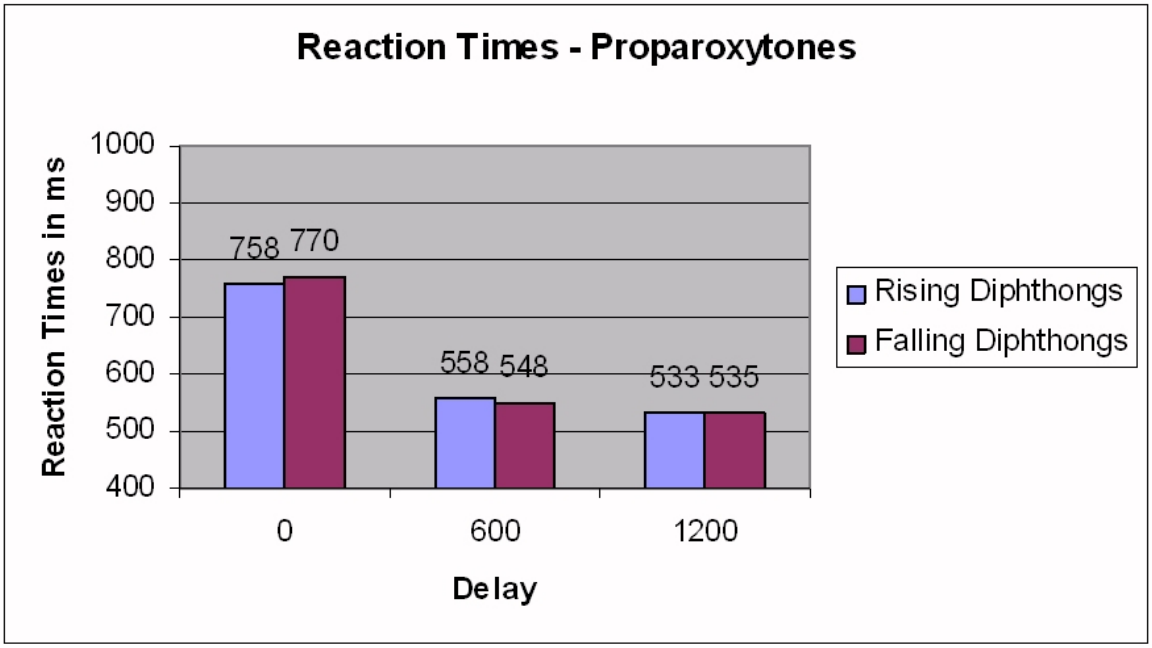

Figure **42**: Experiment 4: Reaction Times – Rising Diphthongs vs. Falling Diphthongs – Only proparoxytones across all delays

There is no significant difference between rising and falling diphthong in any of the delay conditions (0 ms = $p>.99$, 600 ms = $p>.99$, 1200 ms = $p=1.00$). There is also no difference between diphthong types in the penultimate condition, as illustrated below.

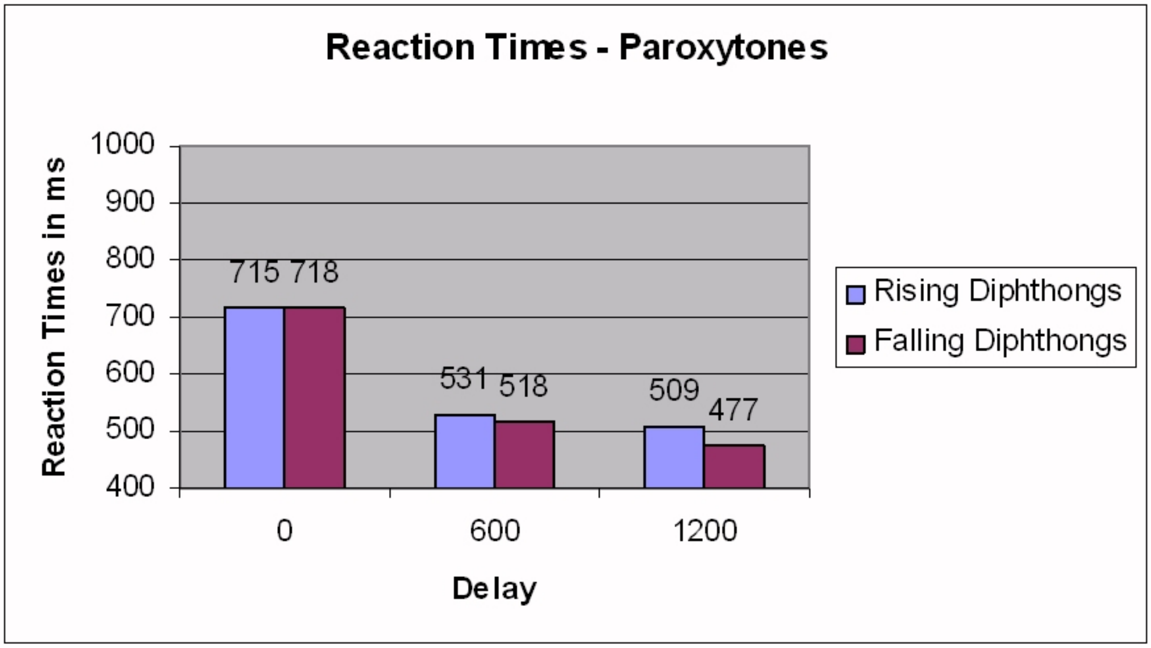

Figure 43: Experiment 4: Reaction Times – Rising Diphthongs vs. Falling Diphthongs – Only paroxytones across all delays

Post-hoc analyses reveal no significant differences between rising and falling diphthong in any delay in the penultimate condition (0 ms = p=1.00, 600 ms = p>.99, 1200 ms = p>.65).

Therefore, comparable to Experiments 2 and 3, Experiment 4 fails to evidence the diphthong effect in the reaction time data. The lack of a statistical difference between the latencies for rising and falling diphthong is likely due to the extremely high error rates across the experiment. The large amount of reaction time data that has been excluded from all reaction time data reduces the statistical power of the variables, attenuating the diphthong effect until it no longer surfaces. As mentioned earlier, it is also possible that latency analyses are an inappropriate tool to test between two proscribed forms. Given the task, accuracy analyses prove to be the more reliable test of phonotactic knowledge and encoding.

Lastly, reaction time data for filler items were also submitted to 2 (frequency) x 3 (stress type) x 3 (delay) repeated measures ANOVAs. The analyses found no significant main effect of frequency [$F1(1, 54) = 0.36$, $MSE = 728$, $p>.55$; $F2(1, 134) = 0.01$, $MSE = 17$, $p>.93$], a significant main effect of stress type [$F1(2, 108) = 9.93$, $MSE = 19545$, $p<.001$; $F2(2, 134) = 3.96$, $MSE = 10528$, $p<.03$], a significant main effect of delay [$F1(2, 108) = 384.11$, $MSE = 3529319$, $p<.001$; $F2(2, 134) = 1259.65$, $MSE = 1413943$, $p<.001$], an interaction of delay and frequency that was significant by participants but not by items [$F1(2, 108) = 13.00$, $MSE = 16907$, $p<.001$; $F2(2, 134) = 2.80$, $MSE = 3138$, $p>.06$], no significant interaction of delay and stress [$F1(4, 216) = 0.90$, $MSE = 1343$, $p>.45$; $F2(4, 134) = 0.91$, $MSE = 1027$, $p>.45$], an interaction of frequency and stress type that was significant by participants but not by items [$F1(2, 108) = 3.88$, $MSE = 6128$, $p<.03$; $F2(2, 134) = 0.92$, $MSE = 2439$, $p>.40$], and no significant interaction of delay, frequency, and stress type [$F1(4, 216) = 1.35$, $MSE = 1714$, $p>.25$; $F2(4, 134) = 1.61$, $MSE = 1804$, $p>.17$]. Interestingly, these ANOVAs indicate that high and low frequency real words were not treated differently in this experiment. The reaction time data for the real word fillers in Experiment 4 are presented in the following figure.

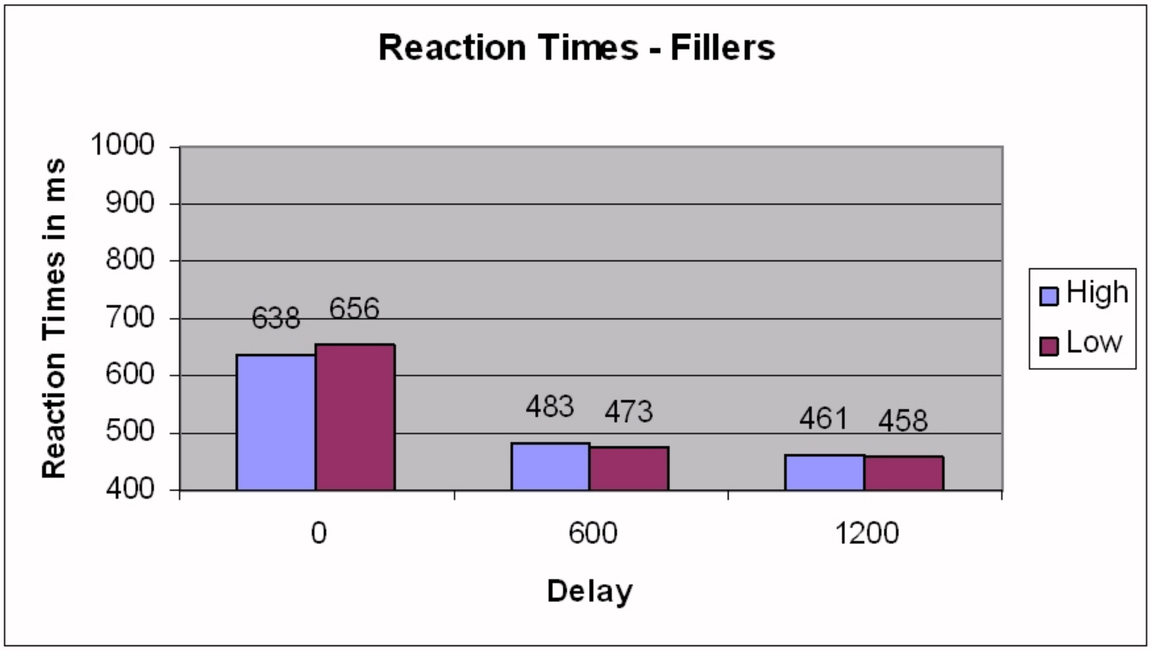

Figure **44**: Experiment 4: Reaction Times – Fillers – Across all delays

Post-hoc tests show no significant difference between high and low frequency items in any of the delay conditions (0 ms = p>.12, 600 ms = p>.68, 1200 ms = p>.99). The data suggest that the delays in Experiment 4 eliminate the frequency effect. The loss of the frequency effect in the delayed naming task is attested in the literature (cf. Janssen, Schirm & Caramazza, submitted). Word frequency is a lexical effect (Bartram, 1976; Jescheniak & Levelt, 1994; Wingfield, 1967), and lexical effects are predicted to disappear in a delayed naming task because the delay conditions allow sufficient time for lexical access, causing any effects produced by this process to vanish (Forster & Chambers, 1973; Monsell et al., 1989, Andrews and Heathcote, 2001). Thus, the loss of the frequency effect is expected, given that 600 and 1200 ms are substantially long delays for real-word items. As is clear from figure 44, 600 ms is already significantly longer than the actual average reaction time for any given item. However, one would predict the frequency effect to surface in the 0 ms delay condition. The following figure illustrates the reaction times for the high and low frequency real words across the three stress conditions in the first delay.

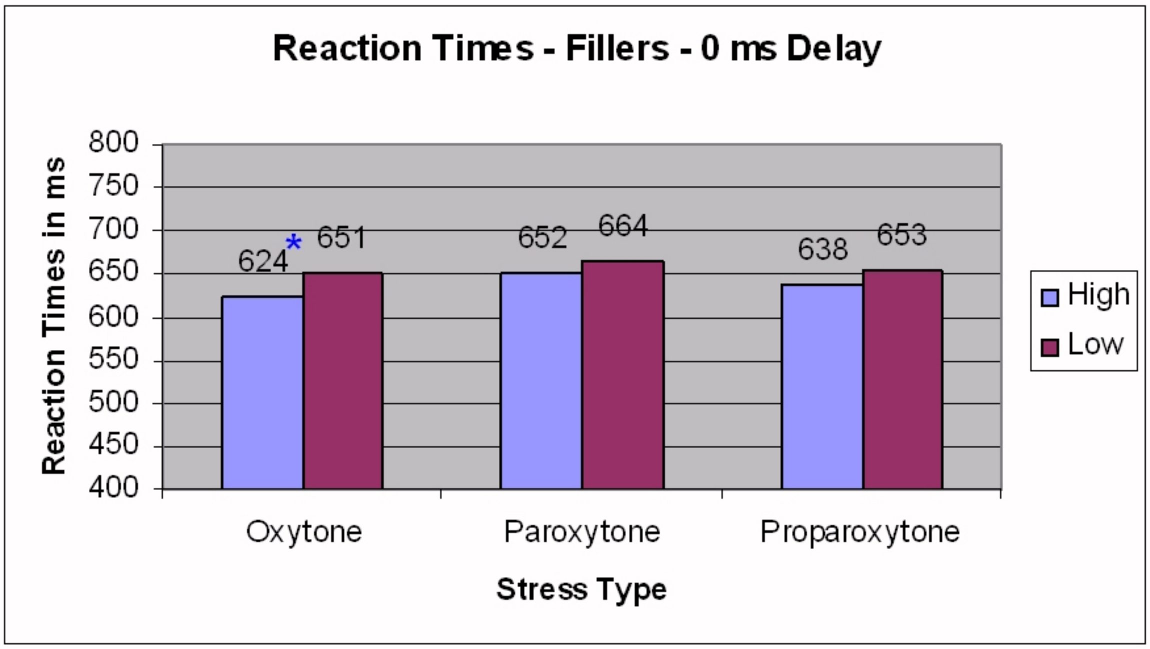

Figure 45: Experiment 4: Reaction Times – Fillers – 0 ms Delay

Descriptively, figure 45 illustrates that lower frequency words were produced more slowly across all stress types in the 0 ms delay. However, this difference is only statistically significant in the oxytone condition ($p<.02$). As argued in the discussion of the accuracy data, the 0 ms delay is an unexpected condition in a delayed naming task. It is also possible here that the noise added by the lack of an actual delay in this condition attenuates the frequency effect. That is to say, descriptively we see a differential pattern between the high and low frequency items. However, the lack of a statistical effect is arguably an artifact of the task conditions. This finding offers support for our interpretation of the accuracy data. The loss of the frequency effect in the 0 ms delay among real words mirrors the loss of the diphthong effect with nonwords. It is apparent from both accuracy and latency data that the lack of a delay in the delayed naming task in Experiment 4 affected participant behavior.

Overall, then, despite the problematic nature of the 0 ms delay, and the lack of sufficient latency data for all predicted effects to be obtained, the findings of Experiment 4 do support the results of Experiments 2 and 3. Diphthongs provoke higher error rates and slower reaction times in the antepenultimate stress condition than do their monophthong control counterparts. This finding offers empirical support based on behavioral data for the quantity sensitive account of the Spanish stress system. The data in these three experiments indicate that native speakers of Spanish process diphthongs and monophthongs differently. In addition, and most importantly, the diphthongs and monophthongs modulate stress placement differently. Diphthongs in the penultimate syllable appear to constrain significantly the possibility of proparoxytone stress. Monophthongs do not show the same pattern.

The subtler diphthong effect was more difficult to interpret within the context of a delayed naming task. Yet, as figure 39 illustrates, there is a clear facilitation of the rising diphthongs 600 ms before any facilitation appears for the falling diphthongs. This finding corroborates the previous results of Experiments 2 and 3. These three experiments indicate that native speakers of Spanish not only represent diphthongs (*heavy syllables*) and monophthongs (*light syllables*) as distinct forms cognitively, but they are also sensitive to the difference between the two diphthong types. The data suggest that falling diphthongs are deemed *heavier* as shown by higher error rates in the 600 ms delay and a later facilitation when viewed across the three delay conditions. Rising diphthongs pattern as *less heavy* in that they present lower error rates than their falling diphthong counterparts in the 600 ms delay. They also evidenced earlier facilitation at the 600 ms delay when viewed across the three delay conditions. Thus, while the delayed naming task requires interpretation across all delay conditions for the effect to be seen, it is apparent that the diphthong effect is indeed a robust finding that has now been replicated twice: with two highly distinct dialects of Spanish, and with two variations on the naming task.

The patterns shown in Experiment 4 are not only interesting from a linguistic perspective. The goal of this experiment, in addition to the replication of the diphthong effect via a different experimental paradigm, was the examination the time course of linguistic processing. Experiment 4 allows for particular insights into the levels of processing in which the encoding of syllable structure and stress assignment become apparent.

Levelt et al. (1999) propose a model of speech production in which syllable structure is encoded online as part of the form level of lexical processing. An opposing view is held by Dell (1986, 1988) in which all syllable encoding is stored with the lexical entry in the mental lexicon. The delayed naming task lends itself well to the comparison of these two models in that the delays are able to isolate effects that originate in the preparation for articulation from effects that emerge from access to the mental lexicon. The reasoning is that any effects that remain after substantial delays in a delayed naming task are attributed to articulation processes rather than lexical access. In this sense, we are able to compare Levelt et al.'s and Dell's models via this experimental paradigm. Specifically, the model proposed by Levelt and his colleagues predicts an effect of syllable structure in Experiment 4 because syllable structure encoding occurs via access to the syllabary, which is posited to be a crucial component of the postlexical form level of speech production. Dell's model would not predict an effect of syllable encoding in Experiment 4, because his account submits that encoding of syllable structure is stored within the lexicon. If a delayed naming task eliminates lexical effects, then there should be no diphthong effect with sufficient time for lexical activation.

Certainly the results from Experiment 4 evidence a diphthong effect across three delay conditions. The facilitation of rising diphthongs after a 600 ms delay and the facilitation of falling diphthongs after 1200 ms suggest online encoding of different syllable structures. This finding is most easily explained via Levelt et al.'s model of speech production. For reference purposes, Levelt et al's (1999) model of speech production is reproduced in the following figure.[29]

---

[29] Figure adapted from Levelt et al. (1999).

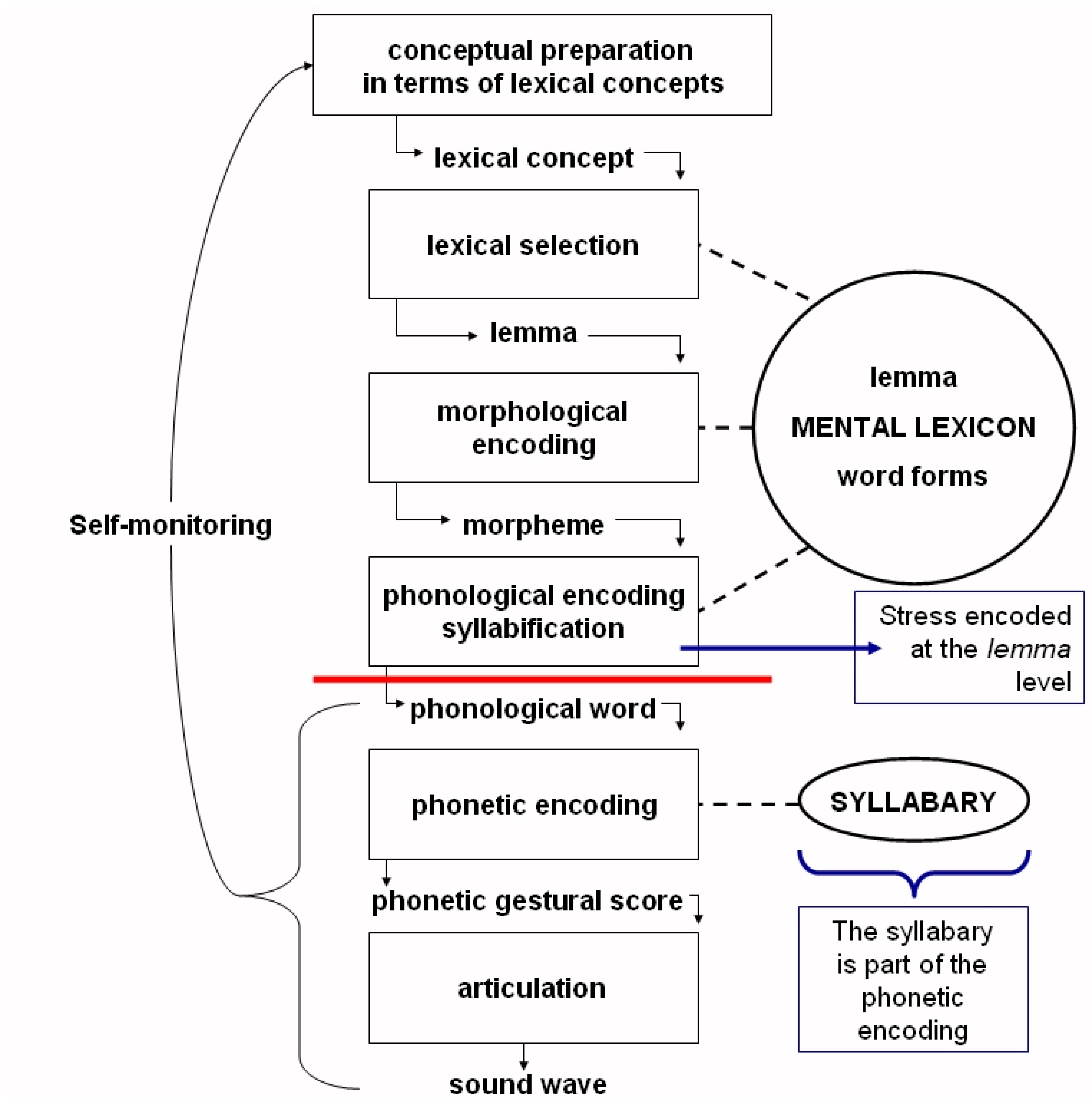

Figure **46**: Levelt et al.'s (1999) model of speech production (2)

As the figure illustrates, syllable structure encoding occurs postlexically as part of the phonetic encoding of the word to be spoken. Given this locus of syllable structure encoding, the model predicts that effects caused by the access to the syllabary should be evidenced in the delayed naming task. It is difficult to account for such an effect in a model such as Dell's in which all knowledge of syllable structure is stored within the lexicon.

However, a crucial aspect of the diphthong effect is the fact that it is precisely an example of how syllable structure constrains stress placement. From a linguist's perspective, this follows traditional models of phonology naturally. Most models of phonology account for stress placement in a bottom-up fashion in which segments combine to form syllables, in which syllables are grouped in to metrical feet that combine to form words, and in which stress is assigned to a particular syllable in the word. The three-syllable window and the restrictions upon it have been precisely delineated in the linguistic literature (see discussion in Chapter 1). The diphthong effect obtained in Experiments 2-4 of this dissertation offers empirical support from a behavioral task that syllable structure indeed constrains stress assignment in Spanish. However, Levelt et al.'s model does not account for the interaction of the encoding of stress and syllable structure information. As depicted in figure 46 above, metrics are encoded at the lemma level. According to the model, metrical encoding consists solely of designating the amount of syllables in the word and assigning stress to one of them. There is no mention of syllable weight or the specific makeup of the syllables themselves. This is considered an artifact of syllabification (Levelt et al., 1999). The syllabary does not come into play until the form level postlexically. Therefore, stress and syllable structure are stored and accessed separately. Given that this model is serial, there is no way for lower levels of processing to interact with levels higher up.

In contrast, Dell's model is highly interactive. There is free interaction between all levels of processing in most connectionist models of speech production. Positing an interaction between metrics and syllable structure causes no problems for a system that is built on a network of highly active connections. If information on both stress placement and syllable structure is stored lexically, the interaction between the two is accounted for within the architecture of the model. However, as discussed previously, it is less clear how Dell's model can explain the differential effect between rising and falling diphthongs over time (i.e. as a function of delay), as evidenced in Experiment 4. In his model, syllable encoding is not incremental. The various syllables in a word are encoded in parallel. Dell's model cannot easily account for the differential treatment of certain syllable types over time.

As a result, both models of speech production offer attractive interpretations of the diphthong effect; although the effect also poses specific problems for each. Levelt et al.'s argument in favor of online encoding of syllables is amenable to the timing of facilitation of accuracy for rising vs. falling diphthongs, as shown in Experiment 4. However, a clearer account of the interaction between stress and syllable structure is essential to explain the diphthong effect. The model proposed by Levelt and his colleagues fails to do so. Dell's model easily rationalizes this interaction but appears to be unable to explain any time course differences between phonotactic patterns in a straightforward manner. Both models would require modifications to their conception of the lexicon and the time course of processing in order to account for the data presented in this dissertation. This topic will be discussed further in Chapter 6.

Chapter 6

Conclusions and Implications

A central goal of this dissertation has been to reexamine via the use of behavioral experimental methods the long-standing question of how best to model native speaker knowledge of the Spanish stress system. This work thus adopts an interdisciplinary approach that bridges perspectives from cognitive psychology and experimental phonology. The experiments reported in this dissertation reveal patterns of results that inform theories of syllable weight and phonotactics in Spanish as well as psycholinguistic models of speech production. In this final chapter, I summarize the main findings of this dissertation and discuss the relevance of the results for both the linguistic modeling of grammatical knowledge and for the psycholinguistic modeling of speech production.

6.1 Summary of Findings

Chapter 1 reviews various perspectives on how to best account for the Spanish stress system. Fundamentally, there are two competing arguments in the generative framework: those in favor of and those against treating quantity sensitivity as part of synchronic Spanish stress (cf. Harris, 1983, 1992, 1995; Roca, 1988, 1990, 1991, 1997, 2005; Lipski, 1997). In the experimental realm, a number of studies have tested native speaker intuitions regarding the acceptability of varying stress patterns (e.g. Face, 2000, 2004, 2005; Bárkányi, 2002; Eddington, 2000, 2004) with conclusions that are equally inconclusive. Regardless of their stance on sensitivity to syllable weight, researchers agree that, at least descriptively, primary stress in Spanish nonverbals systematically falls upon one of the last three syllables of the word. Previous research has described in detail various phonotactic patterns that appear to shrink the putative three-syllable window to one of the two final syllables. Building upon the phonotactic constraints delineated in traditional accounts of Spanish grammar, the experiment discussed in Chapter 2 employs a behavioral method from psycholinguistics to test for cognitive evidence of these theoretical distinctions.

Employing a standard speeded naming task, Experiment 1 compares native speakers' reactions to nonce forms exemplifying various phonotactic restrictions to the three-syllable window. Results from the first experiment accord with traditional theories that argue for a weight-sensitive account of Spanish stress. The behavioral data show that theoretically *heavy* syllables, such as closed penults and diphthongs in the penultimate syllable, effectively constrain proparoxytone stress in that they provoke significantly higher error rates than phonologically licit monophthong controls. This finding, in addition to the replication of the well-established frequency effect among the real-word filler stimuli, validates this methodology as a useful tool to gather empirical support in the form of online, behavioral data for theories that previously relied on metalinguistic knowledge or introspection.

However, Experiment 1 also challenges traditional accounts. The finely-grained nature of the behavioral task is able to identify patterns in the data that have not been identified in previous studies. Specifically, Experiment 1 finds a three-way split between that which we characterize as a synchronic proscription, i.e. proparoxytones with a closed penult (e.g. *dóvalda*); that which is unattested in the lexicon for diachronic reasons, i.e. proparoxytones with palatal final onsets (e.g. *dóvaña*); and phonologically licit gaps, i.e. CV.CV.CV controls (e.g. *dóvasa*). This finding is compelling in that it is indicative of *gradience* within the phonological grammar. Traditional approaches to syllable weight provide a binary classification of syllables: *heavy* and *light*. Results from Experiment 1 suggest that native speakers represent much subtler patterns of syllable-based phonotactic knowledge than previously postulated in treatments of Spanish stress. The findings from Experiment 1 advance the conclusion that we are able to track patterns across our lexicon at various levels of generalization.

In Chapter 3, Experiment 2 explores another subtle effect that traditional grammars cannot easily characterize. The binary nature of traditional theories of syllable weight in Spanish characterizes both rising and falling diphthongs as *heavy*. The theory predicts that two heavy syllables will constrain proparoxytone stress equally. However, as discussed in Chapter 1, rising and falling diphthongs do not pattern identically across the lexicon in that they vary by syllable position. In particular, rising diphthongs in the final syllable regularly fail to attract stress. By contrast, falling diphthongs systematically do attract stress in this position. Experiment 2 builds upon the descriptively substantiated, but statistically deficient, finding of a differential treatment of diphthong types in Experiment 1. The second experiment focuses specifically on rising diphthongs, falling diphthongs, and segmentally-controlled monophthongs. Utilizing this more concentrated experimental design, including twice as many stimuli of each diphthong type, Experiment 2 reveals a significant diphthong effect. As with the three-way split of *proscriptions→diachronic gaps→fully licit gaps* evidenced in Experiment 1, the diphthong effect from Experiment 2 argues for a gradient conception of syllable weight in which falling diphthongs are represented cognitively as the *heaviest* form. Rising diphthongs, which do provoke high amounts of error, yet crucially *significantly fewer* than the falling diphthongs, must therefore be represented in the grammar as *less heavy*. The monophthong controls represent the *lightest* syllable type. This second three-way split, *falling diphthongs→rising diphthongs→monophthongs*, illustrates an even subtler division within the proscription category of the *proscription→diachronic gaps→fully licit gaps* sequence. Therefore, it is clear that the traditional view of binary divisions in syllable weight provides an insufficient and overly narrow perspective on what is actually a delicate, finely-grained, and remarkably subtle continuum consisting of the incorporation of higher level prosodic information, e.g. stress placement, with lower level encoding, e.g. syllable and segmental composition.

The subtlety of the diphthong effect evidenced in Experiment 2 is a surprising result. While its presence is robustly manifest in the accuracy analyses from Experiment 2, it is nonetheless a perplexing and troubling finding, from the perspective of traditional accounts of Spanish stress. Therefore, the goal of Experiment 3 was to replicate the results of Experiment 2 with a different population. One might argue that the effects obtained in the first two experiments are characteristic only of the dialect spoken in Jaén, Spain, where the first two experiments were run. To begin to generalize the findings across speakers of Spanish, regardless of dialect, it is necessary to test a separate group of Spanish speakers. The participants in Experiment 3 were native speakers of a dialect of Spanish spoken in central Mexico. The surface-level differences between the two dialects are numerous (cf. Zamora Vicente, 1960; Penny, 2000; Canfield, 1981; Lope Blanch, 1990; Lipski, 1994). Therefore, a replication of the results from Experiment 2, in spite of the identified phonological divergences between the two speech varieties, would suggest a deeper level of shared knowledge among speakers of Spanish. Indeed, the diphthong effect obtained in Experiment 2 also surfaced in Experiment 3. These results underscore the robust nature and collective quality of the gradient knowledge shared by native speakers of Spanish.

Chapter 5 tests the time course of phonological encoding in the context of two competing models of speech production (Levelt et al., 1999; Dell, 1986, 1988), which differ substantially with respect to the way in which they model syllable-level phonological encoding. As discussed in detail in Chapter 1, Levelt et al.'s model posits that syllables are encoded *after* a lexical item has been accessed in the mental lexicon. High frequency syllables are stored in a mental syllabary that is accessed as part of the postlexical phonological/phonetic encoding process. Dell's model argues, in contrast, that information about the syllables in a word is stored with the lexical item in the lexicon. Given that the diphthong effect is a reflection of syllable structure and how it modulates stress placement, this effect is a promising tool to inform the timing of syllable preparation in speech production.

Experiment 4 employs the delayed naming task to compare the differences in processing between rising and falling diphthongs across time. Differential results for rising and falling diphthongs across delay conditions would suggest a separate locus for syllable-level encoding. This would favor the model proposed by Levelt and his colleagues. If the diphthong effect surfaces without any variation across the delay conditions, this would support Dell's conception of the lexicon in which syllable encoding is a component of the lexical entry. Results from the delayed naming task show that rising and falling diphthongs are differentially facilitated across the delay conditions. The data show that a 600 ms delay is a sufficient amount of time to facilitate significantly the encoding of rising diphthongs, as evidenced by improved accuracy in the second delay condition. However, falling diphthongs require a 1200 ms delay to evidence a significant level of improvement. This different pattern for rising and falling diphthongs across the delay conditions supports the conception of a separate locus for syllable-level encoding.

Levelt et al.'s account of speech production explains this finding best to the extent that the delayed naming task is cited in the literature as a tool that is typically employed to differentiate between lexical and postlexical effects (Forster

& Chambers, 1973; Monsell et al., 1989; Andrews & Heathcote, 2001). Given that the diphthong effect surfaced within the delayed naming paradigm in Experiment 4, we may attribute the finding to the postlexical phonological/phonetic level of encoding. Levelt et al.'s model holds that the syllable is encoded postlexically prior to articulation. It is difficult to explain the time course results of Experiment 4 within Dell's framework in which syllable encoding is stored with the word in the lexicon.

Dell's model, however, captures better the interactive nature of the diphthong effect. As discussed throughout this dissertation, the diphthong effect is a direct result of syllable structure conditioning stress placement. The interaction between these two levels of processing is intrinsic to a connectionist conception of the lexicon, as in Dell's model. Levelt et al.'s model is less congenial to this type of interaction. The data from this dissertation raise serious challenges for the relationship between the stress and syllable levels of encoding as the serialist modelers describe them. In this model, metrical information, e.g. stress, is encoded at the lemma level of the process. It is therefore one of many lexical processes (e.g. morphological encoding, verb subcategorization, etc.). Syllable-level encoding, as mentioned previously, is a postlexical process according to the model. Given that the model is feedforward-only, it is impossible for syllable structure to modulate stress assignment. Such a process would imply that information that is processed postlexically may constrain processing higher up in the model. This is not possible within the framework of Levelt, Roelofs, & Meyer (1999).

Therefore, it is clear from the data that psycholinguistic models must account for two findings in Experiment 4. First, if we assume that the delayed naming task serves the function of demarcating lexical and postlexical effects, then any model of speech production must account for the time course results from Chapter 5 which point to differential patterns for rising and falling diphthongs. Secondly, the tight connection between syllable structure and stress placement needs to be explored more fully. Neither of the prominent models of speech production discussed here (Levelt et al., 1999; Dell, 1986, 1988) is capable of explaining both results.

6.2 Linguistic Import

The results of the four experiments discussed in this dissertation fall in line with the predictions of a weight-sensitive account of the Spanish stress system. While there is much debate in the literature, both synchronically and diachronically, as to whether Spanish is, or was, quantity sensitive, the data in Chapters 2-5 suggest that native speakers cue into patterns across the lexicon in order to make generalizations about the varying behavior of different phonotactic sequences. Those patterns that have been characterized in previous studies as *heavy* do indeed pattern in these data as statistically different from those deemed *light*. And the heavy syllables are those that evidence higher rates of error, and sometimes slower latencies, in the theoretically proscribed conditions. Therefore, the findings of this dissertation could be interpreted as empirical support for theories of Spanish stress that argue in favor of weight sensitivity, or, at the very least, as confirming evidence for the relevance of sensitivity to patterns of syllable structure and its interaction with stress assignment in the synchronic phonology of Spanish.

However, when examined as a whole, the data in this dissertation also argue strongly for a stochastic conception of the mental lexicon. The results are indicative of gradience in the phonological system. Traditional accounts of Spanish stress theory that argue for a sharp division between *light* and *heavy* syllables maintain too broad a perspective to be able to account for the subtler findings obtained in the studies here. It appears rather that native speakers cognitively represent a continuum of acceptability for various phonotactic patterns. Some are treated as "better" or "worse" than others, with putative binary distinctions revealed to be less categorical than traditionally assumed under appropriate experimental conditions.

While traditional terminology, such as "syllable weight" and "quantity sensitivity", is ubiquitous in this thesis, such terms are intended to identify variation in sensitivity to patterns across the lexicon. It is not the aim of this work to operationalize the observed gradience in sensitivity, referred to in the dissertation as "heavy", "less heavy", and "light", as 2, 1.5, or 1 moras, as weight is often characterized in traditional frameworks. Rather, the data speak to a granularity in the system that is better accounted for through multiple levels of generalization that are developed upon observations of the probability of distinct phonotactic sequences in the lexicon. This is not to say that phonologists should do away completely with abstract notions such as the mora. It is clear that a certain level of abstraction is useful and necessary. However, the findings of this dissertation suggest further levels of subtle patterns to which native speakers are highly attuned.

Broadly speaking, I argue that the gradience in acceptability is best understood within the framework of a stochastic or emergentist perspective of grammar. As opposed to traditional models in which phonological knowledge is represented at an abstract and fundamentally categorical level, probabilistic or stochastic models hold that phonology is derived through sensitivity to patterns across the lexicon. We build our phonological grammars, not by excluding all redundant information in order to store purely idiosyncratic information, but rather by cueing into the patterns we perceive across multiple stored tokens of individual speech (cf. Bybee, 2001; Pierrehumbert, 2001b). Within this framework, there is no reason to posit strict categorical distinctions among the generalizable patterns. Because individual speech is highly variable, language is naturally gradient. Thus, memory tokens of such speech will also vary. Therefore, it is logical to conclude that the generalizations we make based on this input may also be of a gradient nature.

The gradience in the system leads us to predict that there should be varying degrees of acceptability across nonce word forms such as those used in the experiments here. Since nonwords are not themselves stored in the lexicon, our reaction to them must be based on their adherence to or their inconsistency with generalized patterns that we have perceived across the actual forms we have stored in our lexicon. This view of the lexicon best explains the results discussed in Experiments 1-4 of this dissertation, in which participants' reactions to different phonotactic patterns exhibit a continuum of acceptability, arguably mirroring the gradience in the phonological system.

6.3 Psycholinguistic Import

The initial purpose of this dissertation was not to *test* the two models of speech production discussed in Chapters 1 and 4. The experiments were not designed with that intent. Rather, they were developed to test in a behavioral manner native speakers' knowledge of and reactions to various constraints on antepenultimate stress. Nevertheless, the results prove informative to psycholinguistic models of language processing. While the findings of these four experiments neither support nor refute either model in particular, they do challenge both models to extend their conception of the lexicon. More specifically, the models are encouraged to specify more clearly the processing of phonology as concerns stress and syllable structure, and the output of these two levels of encoding to phonetic/articulatory levels.

Experimentation probing for differences in treatment between the rising and falling diphthongs in Spanish is a particularly useful methodology in that any effects found must be explained with respect to: 1) where the syllable is encoded, as the difference between the two diphthong types is intrinsically a syllable-level effect, and 2) how syllable encoding interacts with the processing of stress. While both Levelt et al.'s and Dell's models postulate that information about stress placement for words is stored in the lexicon, the former posits postlexical encoding of syllable structure via a mental syllabary, and the latter maintains that all information is stored lexically. The nature of the rising vs. falling diphthong effect allows for a basis for comparison between the two models of speech production.

The findings of this dissertation challenge both models to explain more clearly the interaction between syllable structure and stress. Neither model is capable of accounting in a straightforward manner for both the differential treatment of rising and falling diphthong stimuli over the course of time, which is indicative of a separate, postlexical locus of syllable level encoding, and the means by which syllable structure is able to constrain stress placement in Spanish words. Levelt et al.'s model better accounts for the time course of syllable encoding by positing a separate locus for syllable-level processes. However, this model cannot easily account for the interaction of syllable structure and stress. Dell's model can account for the interaction of syllable structure and stress, via free spreading of activation across a connectionist network where all encoding is stored with the lexical item. However, his model does not clearly explain the apparent postlexicality of the diphthong effect. The data submitted in this dissertation suggest the need for closer evaluation of the loci of prosodic-level encoding and syllabic encoding in either model.

I have argued, both through my review of the linguistic literature and my interpretation of the results of these experiments, for a stochastic, emergentist conception of the lexicon. Strictly localist models of speech production are not amenable to this perspective of the lexicon. Both Levelt et al. (1999) and Dell (1986, 1988) view the lexicon as a “mental dictionary.” Each word has its own listing. Each listing consists of the information necessary for the word’s production and use in speech (morphosyntactic information, metrics, phonemes, etc.). In this sense, both models follow the traditional view of the lexicon originally put forth by linguists in which our knowledge of words consists of an individual abstract representation. If we assume, by contrast, that our knowledge of the words in our language consists not of purely idiosyncratic information, but rather multiple levels of generalization across numerous stored tokens of speech, then the knowledge we represent and access in speech processing is far more detailed than Levelt et al. and Dell put forth in their models.

The concept of a higher distribution of linguistic knowledge across a neural network is not a new concept in the psycholinguistic literature (e.g. Plaut, McClelland, Seidenberg & Patterson, 1996; Seidenberg & McClelland, 1989). While it is reasonable to assume there must be a certain level of generalization at which word shapes are stored, such as those posited by localist models, it is important to take into account that the lexicon consists of much *more* than just one level of generalization. The current construct of the *lemma* in psycholinguistic models is, arguably, too restrictive to account for the granularity of phonological knowledge in the system. As discussed in Chapter 2, the data in Experiment 1 suggest a fine-grained continuum of sensitivity to phonotactic sequences. The differential behavior of native speakers to various levels of proscriptions and gaps is indicative of their ability to track subtle differences in phonotactic patterns and to represent this knowledge cognitively. Psycholinguistic models of speech production need to account for varying levels of phonotactic awareness. Native speakers cognitively represent knowledge of word forms. That is to say, we know the words of our language, how they are syllabified, what phonemes are associated with them, which syllable is stressed, and so forth. However, just as importantly, we *also* are sensitive to higher levels of generalization *across* the lexicon. This sensitivity consists of an acute cognizance of what does not and cannot exist in our language, as well as what forms can and do form our lexicon. Localist models of speech production, therefore, must account for the representation and access of this type of information in the preparation of linguistic forms for speech as well.

Research on phonotactic probability has been able to tease apart the difference between lexical and sublexical knowledge (Vitevitch, 1997, 2002, 2003; Vitevitch et al., 2004, Vitevitch & Luce, 1998, 2005). Vitevitch and colleagues have shown that neighborhood density and phonotactic probability affect different levels of encoding. Specifically, neighborhood density effects arise through competition with other similar lexical items (i.e. similar words in the lexicon compete for activation). Phonotactic probability, on the other hand, is argued to be a *sublexical* effect which arises due to the frequency with which certain sequences of phonological segments appear in words. Higher frequency items are facilitated. This sensitivity to patterns of frequency among phonotactic sequences is stored within the lexicon. If we adapt Vitevitch's arguments to the models of speech production discussed here, one would argue that the "word level" of activation must consist of more than simply information stored with the word entry. Rather, we must assume that access to the mental lexicon entails activating knowledge of a lexical nature, i.e. the encoding of information particular to the word we want to speak, as well as knowledge that Vitevitch and colleagues refer to as sublexical, i.e. knowledge of the combination of sounds across all words in the language.

While Vitevitch and colleagues have studied the difference between lexical and sublexical effects through the careful control of neighborhood density and phonotactic probability, this distinction may also be borne out in the production of nonwords. Nonwords are not stored in the lexicon, yet we are still able to pronounce them. This is because we are aware of the possible sound combinations in our language beyond the individual words themselves. Previous research has shown that speakers access the lexicon when reading nonce forms aloud (e.g. Coltheart et al., 2001; Plaut, McClelland, Seidenberg & Patterson, 1996; Seidenberg & McClelland, 1989; Seidenberg et al., 1996; McCann & Besner, 1987; McCann, Besner & Davelaar, 1988; Taft & Russell, 1992). Therefore, the use of nonwords, such as those employed in the experiments in this dissertation, is a valid means of investigating lexical processing. If we assume multiple levels of phonological knowledge in the lexicon, it is clear that nonwords are still susceptible to effects caused by their phonotactic composition.

It is the challenge of models of speech production, such as Levelt et al. (1999) and Dell (1986, 1988), to account for the gradient nature of phonotactic knowledge discussed in the linguistic literature and evidenced in the experiments in this dissertation. It is clear that native speakers store knowledge of phonotactic patterns in the lexicon. It is not possible for the tightly knit interaction between segments, syllables, and stress to be solely an emergent effect of articulation. Rather, native speakers exhibit a granularity in the acceptability of various phonotactic patterns. This is a result of their susceptibility to knowledge across the lexicon. The "word level" of speech production models needs to grow to encompass much more than word entries in a mental dictionary.

This is not to argue that the concept of a syllabary is an unmotivated construct. It still appears to be the best account for the time course of syllable encoding evidenced in Experiment 4. However, a purely postlexical account of syllable structure is insufficient to account for the data in this dissertation and does not accord with a stochastic conception of the lexicon. It is the challenge of speech production modelers to account for the phonological and phonetic encoding of syllable structure and stress both within the lexicon and across time during production.

6.4 Future implications

This dissertation offers to the fields of experimental phonology and cognitive psychology a research strategy whose interdisciplinary approach has proven useful to inform both linguistic theories of syllable weight and stress placement in Spanish as well as psycholinguistic models of speech production.

The result of the marriage of these two approaches provides a tool that is capable of identifying subtle patterns that had not been attested empirically via the collection of behavioral data. The methodologies borrowed from psycholinguistic research offer to the field of linguistics novel approaches to support or refute theories through the collection and analysis of online data. Behavioral methods, such as those employed in this dissertation, add to the set of tools with which linguists can explore the complex issue of how to model what speakers implicitly know about their languages. Methodologies such as the naming tasks utilized here offer an advantage in that they do not resort to tasks that ask speakers to tap into explicit knowledge (i.e. introspection) and as such have the capacity to tap more directly implicit speaker knowledge.

Questions generated through linguistic theorizing allow psycholinguists to reexamine their conception of the lexicon, as regards the cognitive representation of linguistic knowledge and the processes undertaken in speech recognition, production, and reading. This dissertation, more specifically, focuses on the importance of the syllable. Models of speech production have often focused more closely on larger-scope issues, such as the interaction of concepts and word selection, or the interaction of semantics and phonology. The details of phonological encoding have received less attention. It is clear, however, that the studies discussed in this thesis support previous work (e.g. Carreiras & Perea, 2002; Carreiras, Vergara & Barber, 2005; Álvarez, Carreiras & de Vega, 2000; Álvarez, Carreiras & Taft, 2001; Álvarez, de Vega & Carreiras, 1998) that the syllable is a crucial level of encoding for the Spanish language. Further research examining the interaction of the syllable with other levels of phonotactic patterning will continue to inform language scientists, linguists and psycholinguists alike.

References

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.

Álvarez, C. J., Carreiras, M., & de Vega, M. (2000). Syllable-frequency effect in visual word recognition: Evidence of a sequential-type processing. *Psicológica, 21*, 341-374.

Ávarez, C.J., Carreiras, M., & Perea, M. (2004). Are syllables phonological units in visual word recognition? *Language and Cognitive Processes, 19 (3),* 427-452.

Álvarez, C. J., Carreiras, M. & Taft, M. (2001). Syllables and morphemes: Contrasting frequency effects in Spanish. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*, 545-555.

Álvarez, C.J., de Vega, M., & Carreiras, M. (1998). La sílaba como unidad de activación léxica en la lectura de palabras trisílabas. *Psicothema, 10*, 371-386.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition, 10,* 565-575.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4 (4),* 439-461.

Andrews, S. & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27 (2),* 514-544.

Archangeli, D., & Pulleyblank, D. (1994). *Grounded phonology*. Cambridge, MA: MIT Press.

Aske, J. (1990). Disembodied rules versus patterns in the lexicon: Testing the psychological reality of Spanish stress rules. In K. Hall, J. Koenig, M. Meacham, S. Reiman, & L. Sutton (Eds.), *Proceedings of the sixteenth annual meeting of the Berkeley Linguistics Society* (pp. 30-45). Berkeley: Berkeley Linguistics Society.

Baars, B. J., Motley, J. T., & MacKay, D. (1975). Output editing for lexical status from artificially elicited slips of tongue. *Journal of Verbal Learning and Verbal Behavior, 14,* 382-391.

Bárkányi, Z. (2002). A fresh look at quantity sensitivity in Spanish. *Linguistics, 40 (2),* 375-394.

Bartram, D. J. (1976). Levels of coding in picture-picture comparison tasks. *Memory & Cognition, 4,* 593-602.

Blevins, J. (2003). The independent nature of phonotactic constraints: An alternative to syllable-based approaches. In C. Féry & R. van de Weijer (Eds.), *The syllable in optimality theory*. Cambridge: Cambridge University Press.

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

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Canfield, D. L. (1981). *Spanish pronunciation in the Americas*. Chicago: University of Chicago Press.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27 (6),* 1430-1450.

Caramazza, A., & Yeni-Komshian, G.H. (1974). Voice onset time in two French dialects. *Journal of Phonetics, 2,* 239-245.

Carreiras, M. & Perea, M. (2002). Masked priming effects with syllabic neighbors in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1228-1242.

Carreiras, M. & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain and Language, 90,* 393-400.

Carreiras, M. Vergara, M., & Barber, H. (2005). Early event-related potential effects of syllabic processing during visual word recognition. *Journal of Cognitive Neuroscience (17:11),* 1803-1817.

CELEX Lexical Database (2001). http://www.ru.nl/celex/. Centre for Lexical Information, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Chateau, D., & Jared, D. (2003). Spelling-sound consistency effects in disyllabic word naming. *Journal of Memory and Language, 48*, 255-280.

Cholin, J., Schiller, N. O., & Levelt, W. J. M. (2004). The preparation of syllables in speech production. *Journal of Memory and Language, 50,* 47-61.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Chomsky, N., & Lasnik, H. (1995). The theory of principles and parameters. In N. Chomsky, *The minimalist program* (pp. 13-128). Cambridge, MA: MIT Press.

Clements, G., & Keyser S. (1983). *CV phonology: a generative theory of the syllable*. Cambridge, MA: MIT Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589-608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204-256.

Contreras, H. (1977). Spanish epenthesis and stress. *Working Papers in Linguistics 3,* 9-33. Seattle: University of Washington Press.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283-321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language, 27,* 124-142.

Dell, G. S., Lawler, E. N., Harris, H. D., & Gordon, J. K. (2004). Models of errors and omission in aphasic naming. *Cognitive Neuropsychology, 21 (2/3/4),* 125-145.

Dell, G.S., Schwartz, M.F., Martin, N., Saffran, E.M., & Gagnon, D.A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104 (4),* 801-838.

Den Os, E., & Kager, R. (1986). Extrametricality and stress in Spanish and Italian. *Lingua, 69,* 23-48.

Eddington, D. (2000). Spanish stress assignment within the analogical modeling of language. *Language, 76 (1),* 92-109.

Eddington, D. (2004). A computational approach to resolving certain issues in Spanish stress placement. In T. Face (Ed.), *Laboratory Approaches to Spanish Phonology* (pp.95-115). Berlin: Mouton de Gruyter.

Face, T. L. (2000). The role of syllable weight in the perception of Spanish stress. In H. Campos, E. Herberger, A. Morales-Front & T. J. Walsh (Eds.), *Hispanic linguistics at the turn of the millennium* (pp. 1-13). Somerville, MA: Cascadilla Press.

Face, T. L. (2004). Perceiving what isn't there: Non-acoustic cues for perceiving Spanish stress. In T. L. Face (Ed.), *Laboratory approaches to Spanish phonology* (pp. 117-141). Berlin: Mouton de Gruyter.

Face, T. L. (2005). Syllable weight and the perception of Spanish stress placement by second language learners. *Journal of Language and Learning, 3 (1),* 90-103.

Ferrand, L., Segui, J. & Grainger, J. (1996). Masked priming of word and picture naming: The role of syllabic units. *Journal of Memory and Language (35)*, 708-723.

Flege, J. E., & Hillenbrand, J. (1986). Differential use of temporal cues to the /s/-/z/ contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America, 79 (2)*, 508-517.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627-635.

Frisch, S., Pierrehumbert, J., & Broe, M. (2004). Similarity avoidance and the OCP. *Natural Language & Linguistic Theory, 22,* 179-228.

Frisch, S., & Zawaydeh, B. (2001). The psychological reality of OCP-Place in Arabic. *Language, 42,* 481-496.

Gerfen, C. (2001). A critical view of licensing by cue: Codas and obstruents in Eastern Andalusian Spanish. In L. Lombardi (Ed.), *Segmental phonology in optimality theory: Constraints and representations* (pp. 183-205). Cambridge: Cambridge University Press.

Goldrick, M. (2004). Phonological features and phonotactic constraints in speech production. *Journal of Memory and Language, 51*, 586-603.

Goldrick, M. (in press). Connectionist principles in theories of speech production. In M.G. Gaskell (Ed.), *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press.

Gordon, M. (2004). Syllable Weight. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetic Bases for Phonological Markedness* (pp. 277-312). Cambridge: Cambridge University Press.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674-691.

Griffen, Z. M. & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313-338.

Guitart, J. M. (1997). Variability, multilectalism, and the organization of phonology in the Caribbean Spanish dialects. In F. Martínez-Gil & A. Morales-Front (Eds.), *Issues in the phonology and morphology of the major Iberian languages* (pp. 515-536). Washington, D.C.: Georgetown University Press.

Harris, J. W. (1983). *Syllable structure and stress in Spanish: A nonlinear analysis*. Cambridge, MA: MIT Press.

Harris, J. W. (1992). *Spanish stress: The extrametricality issue*. Bloomington, IN: Indiana University Linguistics Club.

Harris, J. W. (1995). Projection and edge marking in the computation of stress in Spanish. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 867-887). Cambridge, MA: Blackwell.

Hay, J., Pierrehumbert, J., & Beckman, M. (2004). Speech Perception, Well-formedness and the Statistics of the Lexicon. In J. Local, R. Ogden, & R. Temple (Eds.), *Phonetic interpretation: Papers in laboratory phonology VI* (pp. 58-74). Cambridge: University Press.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry, 20,* 253-306.

Hayes, B., Kirchner, R., & Steriade, D. (2004). *Phonetically based phonology*. Cambridge: Cambridge University Press.

Hooper, J., & Terrell, T. (1976). Stress assignment in Spanish: A natural generative analysis. *Glossa, 10,* 64-110.

Hualde, J. I. (2005). *The Sounds of Spanish*. Cambridge: Cambridge University Press.

Hualde, J. I. (1991). On Spanish syllabification. In H. Campos & F. Martínez-Gil (Eds.), *Current Studies in Spanish Linguistics* (pp.475-493). Washington, D.C.: Georgetown University Press.

Hyman, L. (1985). *A theory of phonological weight*. Dordrecht: Foris.

Janssen, N., Schirm, W. & Caramazza, A. (submitted). The semantic interface effect in the picture-word interference paradigm: Evidence for the response selection hypothesis.

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language, 46*, 723-750.

Jared, D., & Seidenberg, M. S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 92-105.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Language, Memory and Cognition, 20,* 824-843.

Kahn, D. (1976). *Syllable based generalizations in English phonology*. Doctoral dissertation, Massachusetts Institute of Technology.

Kenstowicz, M. (1994). *Phonology in generative grammar*. Cambridge, MA: Blackwell.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Cambridge, MA: Blackwell.

Laefer, C. (1992). Patterns of voicing conditioned vowel duration in French and English. *Journal of Phonetics, 20,* 411-440.

Laganaro, M. & Alario, F. (2006). On the locus of the syllable frequency effect in speech production. *Journal of Memory and Language, 55,* 178-196.

Leben, W. (1973). Suprasegmental Phonology. Doctoral dissertation, MIT. Published 1980, New York: Garland Press.

Lesch, M. F. & Pollatsek, A. (1998). Evidence for the use of assembled phonology in accessing the meaning of printed words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24 (3)*, 573-592.

Levelt, W. J. M. (1989). *Speaking, from intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition, 42,* 1-22.

Levelt, W. J. M. (1993). Timing in speech production iwth special reference to word form encoding. *Annals of the New York Academy of Science, 682,* 283-295.

Levelt, W. J. M. (1994). The skill of speaking. In P. Bertelson, P. Eelen & G. d'Ydewalle (Eds.) *International perspectives on psychological science, Volume 1: Leading themes*. Hillsdale, NJ: Erlbaum.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences, 3 (6),* 223-232.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-38.

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition, 50,* 239-269.

Lipski, J. M. (1994). *Latin American Spanish*. New York: Longman.

Lipski, J. M. (1997). Spanish word stress: The interaction of moras and minimality. In F. Martínez-Gil & A. Morales-Front (Eds.), *Issues in the phonology and morphology of the major Iberian languages* (pp. 559-593). Washington, D.C.: Georgetown University Press.

Lloyd, P. M. (1987). *From Latin to Spanish: Historical phonology and morphology of the Spanish language*. Philadelphia: American Philosophical Society.

Lope Blanch, J. (1990). *Atlas lingüístico de México, vol. I*. Mexico City: Colegio de México – Fondo de Cultura Económica.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 14-24.

McCann, R. S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word-frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance, 13*, 693-706.

Mester, A. (1994). The quantitative trochee in Latin. *Natural Language and Linguistic Theory, 12 (1),* 1-61.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language, 29,* 524-545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30,* 69-89.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43-71.

Núñez Cedeño, R. (1986). La /s/ ultracorrectiva en dominicano y la estructura silábica. *Actas del II Congreso Internacional sobre el Español de América,* 337-347.

Núñez Cedeño, R., & Morales-Front, A. (1999). *Fonología generativa contemporánea de la lengua española*. Washington, D.C.: Georgetown University Press.

Oldfield, R.C. & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17,* 273-281.

Otero, C. (1986). A unified metrical account of Spanish stress. In M. Brame, H. Contreras, & F. J. Newmeyer (Eds.), *A festschrift for Sol Saporta*. Seattle: Noit Amrofer.

Penny, R. (2000). *Variation and change in Spanish*. Cambridge: Cambridge University Press.

Penny, R. (2002). *A history of the Spanish language* (2nd edition). Cambridge: Cambridge University Press.

Perea, M., Rosa, E., & Gómez, C. (2002). Is the go/no-go lexical decision task an alternative to the yes/no lexical decision task? *Memory & Cognition, 30 (1),* 34-45.

Pierrehumbert, J. (1999). What people know about sounds of language. *Studies in Linguistic Sciences, 29 (2),* 111-120.

Pierrehumbert, J. (2001a). Stochastic phonology. *Glot International, 5 (2),* 195-207.

Pierrehumbert, J. (2001b). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hooper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137-157). Amsterdam: Benjamins.

Pierrehumbert, J., Beckman, M. E., & Ladd, D. R. (2001). Conceptual Foundations of Phonology as a Laboratory Science. In N. Burton-Roberts, P. Carr, & G. Docherty (Eds.), *Phonological knowledge* (273-304). Oxford: Oxford University Press.

Piggott, G. (1995). Epenthesis and syllable weight. *Natural Language and Linguistic Theory, 13,* 283-326.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56-115.

Rapp, B. & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107 (3),* 460-499.

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 482-503.

Roca, I. (1988). Theoretical implications of Spanish word stress. *Linguistic Inquiry, 19,* 393-424.

Roca, I. (1990). Diachrony and synchrony in Spanish stress. *Journal of Linguistics, 26,* 133-164.

Roca, I. (1991). Stress and syllables in Spanish. In H. Campos & F. Martínez-Gil (Eds.), *Current studies in Spanish linguistics,* (pp. 599-635). Washington, D.C.: Georgetown University Press.

Roca, I. (1997). On the role of accent in stress systems: Spanish evidence. In F. Martínez-Gil, & A. Morales-Front (Eds.), *Issues in the phonology and morphology of the major Iberian languages* (pp. 619-664). Washington, D.C.: Georgetown University Press.

Roca, I. (2005). Saturation of parameter settings in Spanish stress. *Phonology, 22,* 345-394.

Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). *Psychological Review, 111,* 561-572.

Roelofs, A. & Meyer, A. S. (1998). Metrical structure in planning the production of spoken words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24 (4),* 922-939.

Schiller, N. O., Meyer, A. S., Baayen, R. H., & Levelt, W. J. M. (1996). A Comparison of Lexeme and Speech Syllables in Dutch. *Journal of Quantitative Linguistics, 3 (1),* 8-28.

Seidenberg, M.S. (1985). The time course of information activation and utilization in visual word recognition. In D. Besner, T. G. Waller, & E. M. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 199-252). San Diego, CA: Academic Press.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Seidenberg, M.S., Petersen, A., MacDonald, M. C., & Plaut, D. C. (1996). Pseudohomophone effects and models of word recognition. *Journal of Experimental Psychology: Memory and Cognition, 22*, 48-62.

Seidenberg, M. S., Waters, G.S., Barnes, M. A., & Tanenhaus, M. 1. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23*, 383-404.

Shatzman, K.B., & Schiller, N. O. (2004). The word frequency effect in picture naming: Contrasting two hypotheses using homonym pictures. *Brain and Language, 90,* 160-169.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol 1., pp. 143-186). Hillsdale, NJ: Erlbaum.

Stemberger, J. P. (2004). Neighborhood effects on error rates in speech production. *Brain and Language, 90,* 413-422.

Steriade, D. (1997a). Lexical Conservatism. In The Linguistic Society of Korea (Ed.), *Linguistics in the Morning Calm, Selected Papers from SICOL 1997* (pp. 157-179). Seoul: Hanshin Publishing House.

Steriade, D. (1997b). Phonetics in phonology: The case of laryngeal neutralization. Ms. Los Angeles: UCLA.

Steriade D. (1999a). Lexical Conservatism in French Adjectival Liaison. In J. M. Authier, B. Bullock & L. Reed (Eds.) *Formal Perspectives on Romance Linguistics* (pp. 243-270). Amsterdam: John Benjamins.

Steriade, D. (1999b). Alternatives to syllable-based accounts of consonantal phonotactics. In O. Fujimura, B. Joseph, & B. Palek (Eds.), *Proceedings of the 1998 linguistics and phonetics conference* (pp. 205-245). Prague: Charles University Press.

Taft, M., & Russel, B. (1992). Pseudohomophone naming and the word frequency effect. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 45A*, 51-71.

Treiman, R. (1983). The structure of spoken syllables: Evidence from novel word games. *Cognition, 15,* 49-74.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language, 25,* 476-491.

Treiman, R., Kessler, B., Knewasser, S., Tincoff, R., & Bowman, M. (2000). English speakers' sensitivity to phonotactic patterns. In M. Broe & J. Pierrehumbert (Eds.) *Papers in laboratory phonology V: Acquisition and the lexicon* (pp. 269-282). Cambridge: Cambridge University Press.

Treiman, R., Mullennix, J., Bijelic-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124*, 107-136.

Vigliocco, G., Antonini, T. & Garrett, M.F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science, 8,* 314-317

Vitevitch, M. S. (1997). The neighborhood characteristics of malapropisms. *Language and Speech, 40,* 211-228.

Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 735-747.

Vitevitch, M. S. (2003). The influence of sublexical and lexical representations on the processing of spoken words in English. *Clinical Linguistics and Phonetics, 17,* 487-499.

Vitevitch, M. S., Armbrüster, J., & Chu, S. (2004). Sublexical and lexical representations in speech production: Effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 514-529.

Vitevitch, M. S. & Luce, P. A. (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science, 9,* 325-329.

Vitevich, M. S. & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language, 52,* 193-204.

Wiese, R. (1996). *The Phonology of German*. Oxford: Clarendon Press.

Wingfield, A. (1967). Perceptual and response hierarchies in object identification. *Acta Psychologica, 26,* 216-226.

Zamora Vicente, A. (1960). *Dialectología española*. Madrid: Editorial Gredos.

Appendix A

Experiment 1 Stimuli Set

Closed penults

jómaldo	jomaldo
góvarta	govarta
córanta	coranta
dóvalda	dovalda
póvalpo	povalpo
lóranta	loranta
fárambo	farambo
sórambo	sorambo
rádalda	radalda
fápanta	fapanta
nódasto	nodasto
mómasto	momasto
bánarto	banarto
závarna	zavarna

"n" as final consonant

cónabon	conabon
pómagan	pomagan
góragon	goragon
jóvapon	jovapon
lóvapon	lovapon
póragan	poragan
sóragan	soragan
fádanon	fadanon
báragan	baragan
mónagan	monagan
dólagan	dolagan
náfagon	nafagon
zámacan	zamacan
tánagon	tanagon

Rising diphthong in penult		*Falling diphthong in penult*	
dóbiana	dobiana	góvaino	govaino
lótiago	lotiago	jómaiga	jomaiga
zátiana	zatiana	cótaiga	cotaiga
jóriana	joriana	dótaiga	dotaiga
móriago	moriago	póvaino	povaino
sóriago	soriago	lóraiga	loraiga
fátiago	fatiago	fáraiga	faraiga
báviana	baviana	nóraiga	noraiga
pámiana	pamiana	rámaino	ramaino
gómiana	gomiana	fávaiga	favaiga
ráriago	rariago	bátaiga	bataiga
cóviana	coviana	závaino	zavaino
nátiago	natiago	sómaino	somaino
támiana	tamiana	mómaino	momaino

Palatal final onsets

Singleton		***Digraph***	
dóvaña	dovaña	góvacho	govacho
rádayo	radayo	jónalla	jonalla
cófaña	cofaña	cóvacha	covacha
jóraño	joraño	dóvacho	dovacho
mósaño	mosaño	póvallo	povallo
sávayo	savayo	lóralla	loralla
bávaña	bavaña	fáracho	faracho
lótaño	lotaño	sóvacho	sovacho
zátaña	zataña	rádalla	radalla
fátayo	fatayo	fápalla	fapalla
pádaya	padaya	nódacho	nodacho
gódaya	godaya	mómacha	momacha
náfayo	nafayo	bánallo	banallo
tánaya	tanaya	záralla	zaralla

Monophthong Controls

bánaga	banaga	dóvasa	dovasa
lóraga	loraga	zátana	zatana
sódano	sodano	fátaga	fataga
cótana	cotana	rádano	radano
nóraga	noraga	násago	nasago
pámaca	pamaca	fátago	fatago
góvaro	govaro	bávaza	bavaza
lótago	lotago	sórago	sorago
dófaga	dofaga	córaga	coraga
mósago	mosago	zátano	zatano
támaca	tamaca	mómaso	momaso
jómaga	jomaga	póvapo	povapo
fáraga	faraga	jórava	jorava
rádago	radago	gódana	godana

Fillers – High frequency

Proparoxytones

máquina	lógica	técnica	lágrimas
método	título	género	médico
término	número	línea	crítica
práctica	público	físico	cámara
música			

Paroxytones

derecho	lectura	mercado	vestido
manera	mentira	dinero	fortuna
camisa	figura	ventana	tamaño
gobierno	persona	memoria	modelo
cultura	cadena	futuro	misterio
cocina	caballo	distinto	camino
botella	conjunto	marido	comida
secreto	tristeza	sistema	

Oxytones

natural	calidad	cultural	soledad
popular	profesor	propiedad	militar
cantidad	juventud	personal	director
general	voluntad	superior	capital
principal	nacional	libertad	sociedad
material	familiar		

Fillers – Low frequency

Proparoxytones

trámite	gótico	búfalo	pócima
rábano	tónica	vínculo	náufrago
plátano	vértice	pómulo	válvula
víbora	ráfaga	cólico	cláusula
ciática	vómito	sífilis	píldora
sílaba	dígito	báscula	

Paroxytones

cuchara	paleta	califa	canica
bellota	bisagra	bombero	bautizo
canela	canguro	cazuela	gemelo
bizcocho	bocina	dialecto	coraza
paella	canario	decreto	bayeta
cajero	caldera	botijo	bolero

Oxytones

pesadez	tirador	dejadez	senador
ruiseñor	secador	titular	comprador
coronel	sucursal	tropical	funeral
vestidor	tocador	diagonal	girasol
delantal	tenedor	litoral	marginal
varonil	virginal	solidez	

Appendix B

Experiments 2-4 Stimuli Set

Rising diphthongs

báviana	baviana
cóviana	coviana
dóbiana	dobiana
fátiago	fatiago
páviego	paviego
rámiena	ramiena
sómiena	somiena
táviego	taviego
záriego	zariego
pámiana	pamiana
ráriago	rariago
sóriago	soriago
támiana	tamiana
zátiana	zatiana
jómiego	jomiego
lóriena	loriena
mómiena	momiena
náriena	nariena
gómiana	gomiana
jóriana	joriana
lótiago	lotiago
móriago	moriago
nátiago	natiago
bátiego	botiego
cótiego	cotiego
dótiego	dotiego
fáriena	fariena
góviego	goviego

Falling diphthongs

bátaiga	bataiga	mómaino	momaino
cótaiga	cotaiga	nóraiga	noraiga
dótaiga	dotaiga	póvaino	povaino
fáraiga	faraiga	rámaino	ramaino
báveino	baveino	sómaino	somaino
lóteiga	loteiga	ráreiga	rareiga
móreiga	moreiga	sóreiga	soreiga
náteiga	nateiga	támeino	tameino
pámeino	pameino	záteino	zateino
fávaiga	favaiga		
góvaino	govaino		
jómaiga	jomaiga		
lóraiga	loraiga		
závaino	zavaino		
jóveino	joveino		
cóveino	coveino		
dóveino	doveino		
fáteiga	fateiga		
gómeino	gomeino		

Monophthong Controls

rádano	radano	dóveno	doveno
gómeno	gomeno	nárega	narega
zárego	zarego	córaga	coraga
bávaza	bavaza	pámaca	pamaca
mómena	momena	jómaga	jomaga
fátaga	fataga	bánaga	banaga
góvaro	govaro	mómaso	momaso
dósego	dosego	mósago	mosago
póvapo	povapo	sórago	sorago
lóraga	loraga	jóteno	joteno
sódano	sodano	fátago	fatago
lótago	lotago	nóraga	noraga
jómego	jomego	záteno	zateno
rádago	radago	lóreda	loreda
lórega	lorega	pámeno	pameno
nárena	narena	zátano	zatano
fáraga	faraga	sómena	somena
góvego	govego	sórega	sorega
bánego	banego	fárena	farena
támaca	tamaca	jórava	jorava
cótana	cotana	távego	tavego
pávego	pavego	dófaga	dofaga

Monophthong Controls (cont.)

násago	nasago	rárega	rarega
báveno	baveno	rámena	ramena
cótego	cotego	cóveno	coveno
fátega	fatega	dóvasa	dovasa
mórega	morega	zátana	zatana
támeno	tameno	gódana	godana

Nonword Fillers

pajeto	nalato
zodilo	colupo
tarilo	fofero
fonega	litime
subima	manigo
miresa	parilo
ranaba	gerida
vanezo	suleda
ligoco	pirago
culuma	boleda
tanavo	pomera
jubato	maneco
videna	jariga
pelato	varido

Real Word Fillers – High frequency

Proparoxytones

máquina	lógica	técnica	lágrimas
método	título	género	médico
término	número	línea	crítica
práctica	público	físico	cámara
música			

Paroxytones

derecho	lectura	mercado	vestido
manera	mentira	dinero	fortuna
camisa	figura	ventana	tamaño
gobierno	persona	memoria	modelo
cultura	cadena	futuro	misterio
cocina	caballo	distinto	camino
botella	conjunto	marido	comida
secreto	tristeza	sistema	

Oxytones

natural	calidad	cultural	soledad
popular	profesor	propiedad	militar
cantidad	juventud	personal	director
general	voluntad	superior	capital
principal	nacional	libertad	sociedad
material	familiar		

Real Word Fillers – Low frequency

Proparoxytones

trámite	gótico	búfalo	pócima
rábano	tónica	vínculo	náufrago
plátano	vértice	pómulo	válvula
víbora	ráfaga	cólico	cláusula
ciática	vómito	sífilis	píldora
sílaba	dígito	báscula	

Paroxytones

cuchara	paleta	califa	canica
bellota	bisagra	bombero	bautizo
canela	canguro	cazuela	gemelo
bizcocho	bocina	dialecto	coraza
paella	canario	decreto	bayeta
cajero	caldera	botijo	bolero

Oxytones

pesadez	tirador	dejadez	senador
ruiseñor	secador	titular	comprador
coronel	sucursal	tropical	funeral
vestidor	tocador	diagonal	girasol
delantal	tenedor	litoral	marginal
varonil	virginal	solidez	

Michael Shelton ~ Vita

Education

2007 Ph.D. in Spanish (Hispanic Linguistics), Option in Applied Linguistics, Doctoral Minor: Linguistics, ***The Pennsylvania State University***

2003 M.A. in Spanish, ***The Pennsylvania State University***

2000 B.S. in Spanish, B.S. in German, Minor: Intercultural Communication, ***St. Cloud State University, MN***

Refereed Conference Presentations

2007 Shelton, M., C. Gerfen, & N. Gutiérrez Palma. ***Proscriptions… gaps…and something in between: An experimental examination of Spanish phonotactics***. Paper presented at the 37th Linguistic Symposium on Romance Languages, Pittsburgh, PA.

2005 Gerfen, C., M. Shelton, & N. Gutiérrez Palma. ***Three degrees of weight in Spanish***. Paper presented at the Hispanic Linguistics Symposium, Workshop on Language Processing, University Park, PA.

2005 Shelton, M., C. Gerfen, & N. Gutiérrez Palma. ***Syllable weight and the cognitive representation of rising versus falling diphthongs in Spanish***. Poster presented at the XIV Conference of the European Society for Cognitive Psychology, Leiden, The Netherlands.

Invited Papers and Other Presentations

2006 Shelton, M. ***Aproximaciones experimentales al estudio del proceso de asignación del acento en español***. (Experimental approaches to the study of stress assignment processes in Spanish). Presented at the University of Jaén, November 10, 2006. Jaén, Spain.

2006 Shelton, M., C. Gerfen, & N. Gutiérrez Palma. ***The cognitive reality of quantity sensitivity in Spanish: Rising versus falling diphthongs and the three-syllable window.*** Poster presented at the 21st Annual Graduate Exhibition, The Pennsylvania State University, University Park, PA.

Research Support and Awards

2006-2007 ***Edwin Erle Sparks Fellowship***, The Pennsylvania State University, Semester release from teaching, including tuition and stipend

2006-2007 ***Humanities Initiative Dissertation Fellowship***, The Pennsylvania State University, Semester release from teaching, including tuition and stipend

2005-2006 ***College of Liberal Arts Dissertation Support Grant***, The Pennsylvania State University, $1000, Competitive grant for dissertation research support

2004-2005 ***College of Liberal Arts Dissertation Support Grant***, The Pennsylvania State University, $1500, Competitive grant for dissertation research support

2004-2005 ***Jesús Díaz Award***, The Pennsylvania State University, $200**,** Scholarship for exemplary achievement in the areas of service, teaching, and academics in the Department of Spanish, Italian and Portuguese