Optimal diphthongs : an OT analysis of the acquisition of spanish diphthongs

Alice Krause
University at Albany, State University of New York, alicelkrause@gmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/919

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
Optimal Diphthongs: An OT Analysis of the Acquisition of Spanish Diphthongs

by

Alice Krause

A Dissertation
Submitted to the University at Albany, State University of New York
In Partial Fulfillment of
The Requirements for the Degree of
Doctor of Philosophy

College of Arts & Sciences
Department of Languages, Literatures and Cultures
2013
Optimal Diphthongs: An OT Analysis of the Acquisition of Spanish Diphthongs

by

Alice Krause

COPYRIGHT 2013
Abstract

This dissertation investigates the acquisition of Spanish diphthongs by adult native speakers of English. The following research questions will be addressed:

1) How do adult native speakers of English pronounce sequences of two vowels in their L2 Spanish at different levels of acquisition?

2) Can OT learnability models, specifically the GLA, account for the pronunciation of L2 diphthongs? If so, what constraints do learners use and how do these constraints interact? If not, what other model(s) might offer an improved analysis of L2 diphthongs?

Participants completed two production tasks, a Nonsense Word task and a Question & Answer task. The participants were divided by level of acquisition – Beginner, Intermediate, Advanced – and there was a Native Speaker Control group. After the data was collected, F2 values and duration of vowel sequence were measured and used to categorize the pronunciations as monophthongs, diphthongs, or hiatus. It was found that the use of diphthongization increased with level of acquisition in the data for the Question & Answer task. Data from the Nonsense Word task did not reveal the same pattern; instead, the level of diphthongization was more or less equal across all levels of acquisition and with the Native Speaker Control group. The OT account was able to explain most of the data in this study. The GLA proved successful in demonstrating how constraints interact in the pronunciation of L2 diphthongs. However, there were L2 pronunciations for which OT could not account. It is suggested that linguistic models based on lexical frequency may offer insight into how to account for these pronunciations.
Acknowledgements

Thank you to my dissertation director, Professor Maurice Westmoreland, for having confidence in me that I could do this project and for allowing me freedom in how I worked. Thank you to Professor Lotfi Sayahi for your straightforward guidance and your interest in my project. Thank you to Professor Lee Bickmore for allowing me two independent studies in which I developed the ideas for this project and also for being a great instructor. Thank you to Professor Rose-Marie Weber for making me a better writer. Thank you to Professor Elizabeth Lansing for your support in so many ways.

Thank you to my husband. Thank you for being smart, supportive, patient, and kind.

Thank you to all of my colleagues, past and present. Most of you participated in this project in some way. Quite a few of you have become some of my closest friends.

Thanks for all the memories.
# Table of Contents

**Chapter 1: Introduction**

1.1 Syllables and diphthongs  
1.2 Diphthongs and Second Language Acquisition  
1.3 SLA and OT  
1.4 Research Questions and Overview of the Study  

**Chapter 2: Rules and Predictions**

2.1 Introduction  
2.2 Vowels and Diphthongs in Spanish and English  
2.3 The Syllable in Spanish and English  
2.4 Sonority  
2.5 Syllabic Processes  
2.6 The Phonemic Status of Glides  
2.7 Making Predictions  
2.8 Conclusion  

**Chapter 3: Constraints and Predictions**

3.1 Introduction  
3.2 Optimality Theory  
3.3 Spanish Syllable  
3.4 The English Syllable  
3.5 OT and SLA  
3.6 Predictions  
3.7 Conclusion
7.3 Recommendations for Future Research

Bibliography

Appendix A

Appendix B

Appendix C

Appendix D
Chapter 1: Introduction

1.1 Syllables and Diphthongs

This dissertation investigates the acquisition of Spanish diphthongs by adult native speakers of English. The study of diphthongization allows for various insights into second language (L2) speech. Diphthongization requires knowledge of the vowel system, implicit and/or explicit awareness of phonological processes, and an understanding of syllable structure. Pronouncing optimal diphthongs means that the L2 learner has acquired knowledge in these three areas.

Syllabification and diphthongization are linked in that diphthongization alters syllabic structure. Syllables and diphthongs are similar in that their descriptions in the literature have sometimes been controversial, especially concerning their definitions and defining their sub-components. The existence of the syllable as a linguistic unit in phonological theory is relatively recent, and the parsing debates in some languages, such as English, are ongoing. The definition of a diphthong is tied to the concept of the syllable. Although a diphthong may be readily identifiable, the categorization and phonemic status of its on- or off-glide is not so easily rendered in the literature.

How syllables are defined and conceptualized has changed in the literature over time. Stetson (1951) takes a physiological view of syllables, discussing respiratory muscle pulses and movements of the chest wall. Saussure sees syllables a series of “irreducible units, opening and closing sounds” (1966: 59). Syllabification then is based on a “succession of implosions and explosions” (Saussure 1966: 60). Trubetzkoy defines the syllable, and the nucleus in particular, as “the smallest prosodic unit” and focuses mostly on the nucleus (1969: 182). Fudge (1969) takes an abstract view of the syllable
and proposes a phonetic and a phonemic syllable. Hyman (1985) discusses the syllable in terms of phonological weight. Clements defines the syllable in terms of sonority (1990). Selkirk (1982) argues for the existence of the syllable based on the fact that many phonological rules and phonotactic constraints only function within the domain of the syllable. She also proposes the internal structure of the syllable that is most commonly used, that of the division of the syllable into onset and rime and the further division of the rime into nucleus and coda. In order to make this distinction, Selkirk (1982) relies on Pike’s (1967) division of syllables into a peak and its margins. The syllable has since figured prominently in linguistic and psycholinguistic research, in first and second language acquisition, and in phonological theory. In Optimality Theory (OT) (Prince & Smolensky 1993/2004), much of constraint interaction revolves around the syllable.

Diphthongs are most controversial when it comes to the phonemic status of glides, but the definition of diphthongs also varies in the literature. Lehiste & Peterson define a diphthong as a “vocalic syllable nucleus containing two target positions” (1961: 277). Saussure defines a diphthong as “a special kind of implosive link” (1966: 61). It is “an implosive link in which the second element is relatively open” (Saussure 1966: 61). Trubetzkoy (1969) defines a diphthong as being comprised of two vowel phonemes. Both Saussure (1966) and Trubetzkoy (1969) point to uncertainty in how to transcribe glides in diphthongs. Hyman states, “Perhaps the most problematic segment type for all theories of phonology is the class of glides” (1985: 77). Glides and their corresponding high vowels are generally seen as allophones of /i/ and /u/ that vary based on syllabicity.
This discussion will be taken up in Chapter 2, along with a discussion as to which part of the syllable a glide in a diphthong pertains.

1.2 Diphthongs and Second Language Acquisition

Contrastive analyses have long pointed to vowels as a potential source of difficulty in the L2 Spanish pronunciation of English speakers. Contrastive accounts, such as Stockwell & Bowen (1965), have highlighted the many possible pitfalls for English speakers in their L2 Spanish pronunciation. Later investigations have since offered explanations for first language (L1) transfer that go beyond contrastive analyses (Morrison 2003, 2006; Menke & Face 2010, Lansing 2001, Reeder 1998). These studies find that the L1 does influence L2 vowel pronunciation and that pronunciation approaches the L2 target pronunciation as both exposure to the L2 and L2 proficiency increase.

Stockwell & Bowen (1965) also point to difficulties in L2 diphthong pronunciation by English speakers in L2 Spanish. Subsequent research that goes beyond offering contrastive details has not appeared until recently. Zárate-Sández (2011) investigates L2 learners’ intuitions about syllabification of two vowel sequences, which entails their knowledge about diphthongization. He found that L2 learners across levels of acquisition shared intuitions about diphthongization that differed from those of native speakers and that the L2 learners favored hiatus over diphthongization. Cross-linguistic influence played a significant role in choosing a hiatus over a diphthong.

MacLeod (2012) investigates the production of rising diphthongs in novel words in the L2 Spanish of English speakers. She finds that there was not a difference in duration for diphthongs by any of the participant groups in her study, but that hiatus was
longer for L2 learners than for native speakers of Spanish. She found that Spanish vowels were longer for the L2 learners than for native speakers, hence the longer duration for hiatus. She found that there was an effect for level of acquisition, in that hiatus duration was shorter as the level of acquisition increased.

Carlson (2006, 2010) offers a frequency based investigation of the L2 acquisition of the mid vowel/diphthong alternations in Spanish verbs. He finds that learners’ production does not match that found in the Spanish lexicon, but that learners are sensitive to statistical patterns in language. However, these sensitivities are manifested in different ways for native speakers and L2 learners.

Based on these studies, it is evident that English vowel pronunciation influences intuitions about diphthongization and diphthong production. While the results between participant groups may be seemingly similar, there are distinctions between native speaker and L2 learner groups. These differences have been studied in the field of Second Language Acquisition (SLA) under various theoretical frameworks.

1.3 SLA and OT

The influence of the L1 on L2 production has been a topic of debate in SLA. Varying from the contrastive models in the 1960s to theories of universal markedness from the late 20th century, the L1 is in some way present in the discussion of how adult learners acquire L2 phonologies. As a point of departure from the descriptive work done under the framework of Contrastive Analysis, Tarone (1980) was one of the first to systematically gather performance data from L2 learners. She found that L1 transfer and a preference for universally unmarked structures were both factors affecting interlanguage phonology. Eckman (1981), in an effort to predict difficulty in L2
phonological acquisition, also found that the L1 influences interlanguage processes in acquiring syllable structure. The body of work by Flege and colleagues examines phonetic data to decipher the extent of L1 influence on L2 pronunciation. Flege & Hillenbrand (1984) found that perceptual categories in the L1 influence the L2, so that an L2 sound that does not have a perceptual counterpart in the L1 may be easier to acquire. Flege & Davidian (1984) find that L1 transfer and developmental processes based on universal constraints affect L2 speech production. Additionally, Flege, Frieda, & Nozawa (1997) find that the amount of L1 use influences L2 pronunciation. Archibald agrees with previous findings that L1 transfer influences L2 phonology, but he argues that “transfer is not always simple: complex structures and principles are transferring” (1998: 209). Many of these studies point to L1 transfer, but they also point to the influence of universal principles or constraints and the concept of markedness.

While markedness was mentioned in pre-OT accounts of interlanguage phonology, there was no mechanism to fully incorporate it in explanatory theories (Eckman 2004). OT directly incorporates markedness in a theory of constraint interaction between markedness and faithfulness constraints. OT learnability models rely on the concept of universal markedness to explain language acquisition processes. The Gradual Learning Algorithm (GLA) (Boesrma & Hayes 2001) is an OT learnability model designed to account for L1 acquisition that has since been utilized in SLA. It allows for variation in the output through the use of constraint ranges, so that the constraints overlap to permit changes in constraint ranking along a strict/lax continuum. As learners receive more input they gradually adjust the constraint rankings by demoting
markedness constraints in favor of being faithful to the input. The constraint ranking should, over time, converge on the native speaker constraint ranking.

OT has proven successful in accounting for various aspects of interlanguage phonology. Hancin-Bhatt & Bhatt (1997) demonstrate that OT is able to account for the combined effects of L1 transfer and developmental processes in the acquisition of L2 syllables. Broselow, Chen & Wang (1998) provide evidence for the emergence of the unmarked (McCarthy & Prince 1994) in L2 phonological acquisition and highlight the role that universal markedness plays in SLA. Swanson (2001) presents various stages of constraint ranking using an OT learnability algorithm developed by Tesar & Smolensky (1996), and she questions whether faithfulness constraints can also be demoted. Broselow & Xu (2004) utilize the GLA and find that its predictions based on frequency are not always realized because L2 learners rely on perceived rather than absolute frequency. Escudero & Boersma (2001, 2004; Boersma & Escudero 2004) have shown that OT learnability models, including the GLA, can account for L2 vowel perception in languages such as Dutch or English and Spanish that have different sized vowel inventories. This literature is further reviewed in Chapter 3, and the tenets of the GLA are used to explain variability in output for L2 diphthongization in Spanish.

1.4 Research Questions and Overview of the Study

The research questions for this study are as follows:

1) How do adult native speakers of English pronounce sequences of two vowels in their L2 Spanish at different levels of acquisition?
Can OT learnability models, specifically the GLA, account for the pronunciation of L2 diphthongs? If so, what constraints do learners use and how do these constraints interact? If not, what other model(s) might offer an improved analysis of L2 diphthongs? This study answers these questions in the following chapters.

Chapter 2 compares and contrasts the vowel systems and the syllable structure of Spanish and English. The information presented in this chapter is based on rule-based research in derivational phonology. Chapter 2 also presents pre-OT SLA research, focusing on contrastive accounts and the effects of L1 transfer to make predictions for L2 pronunciation of Spanish diphthongs.

Chapter 3 presents OT analyses of Spanish and English syllabification. In addition, this chapter reviews SLA research done in an OT framework with the purpose of predicting how constraints might interact in the production of L2 diphthongs in Spanish.

Chapter 4 presents the research methodology for this study. Participants completed two production tasks, which were a Nonsense Word task and a Question & Answer task. The participants were divided by level of acquisition – Beginner, Intermediate, Advanced – and there was a Native Speaker Control group. After the data was collected, F2 values and duration of the vowel sequences were measured and used to categorize the pronunciations as monophthongs, diphthongs, or hiatus.

Chapter 5 reveals the study’s results. It was found that the use of diphthongization increased with level of acquisition in the data resulting from the Question & Answer task. The data resulting from the Nonsense Word task did not reveal the same pattern. Instead,
the level of diphthongization was more or less equal across all levels of acquisition and with the Native Speaker Control group.

Chapter 6 offers an OT analysis of the various processes used to pronounce the vowel sequences in this data set. The OT account was able to explain most of the data. The GLA proved successful in demonstrating how constraints interact in the pronunciation of L2 diphthongs. However, there were L2 pronunciations for which OT could not account. It is suggested that linguistic models based on lexical frequency may offer insight into how to account for these pronunciations.

Chapter 7 concludes this study and considers avenues for future research.
Chapter 2: Rules and Predictions

2.1 Introduction

This chapter begins by describing and comparing Spanish and English vowel systems. The quantity and quality of the vowels in the two systems provide much to compare, and the variation between the two systems can present difficulty to English speakers when acquiring Spanish as a second language. Spanish has a simple and cross-linguistically common vowel system. Most dialect variation in the language concerns the pronunciation of consonants. For this reason, a specific dialect of Spanish will not be discussed, as the general focus of this study is vowel and diphthong pronunciation, something that is relatively stable across dialects of Spanish. On the other hand, General American English (GAE) has a more complex vowel system than that of Spanish, and most dialect variation in English is found in vowel pronunciation (Hualde 2005: 120, 124). A rule-based account and a comparison of Spanish and English syllabification and stress patterns will follow the discussion of vowel systems, as they relate to the processes of gliding, diphthongization, and other relevant vowel phenomena in both languages. Predictions will then be made as to which aspects of the Spanish vowel system would likely present difficulty to English speakers acquiring Spanish as a second language. The predictions made in this chapter follow those that are currently present in the literature and that are based on a derivational theory of phonology.

---

1 General American English is the term used to refer to the variety of English spoken in much of the United States, with the exclusion of the southern states and the eastern seaboard. It is the variety most commonly heard in news broadcasts and other media, as it is seen to be void of strong regional characteristics (McMahon 2002: 69).
2.2 Vowels and Diphthongs in Spanish and English

This section will describe and compare the vowel inventories of Spanish and English. The formation of diphthongs in both languages will also be discussed.

2.2.1 Vowels and Diphthongs in Spanish

Spanish has a five vowel system consisting of the phonemes /i/, /e/, /a/, /o/, /u/, as can be seen below in Fig. 2.1. The non-back, unrounded vowels are /i/, /e/, /a/, and the back, rounded vowels are /o/, /u/.

**Fig. 2.1 Spanish vowels**

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>/i/</td>
<td>/u/</td>
</tr>
<tr>
<td></td>
<td>/e/</td>
<td>/o/</td>
</tr>
<tr>
<td>Low</td>
<td>/a/</td>
<td></td>
</tr>
</tbody>
</table>

There is some allophonic variation in the Spanish vowel system. The high vowels /i/, /u/ have as allophones the corresponding front and back glides [i], [u]. As a rule, these glides occur when an unstressed high vowel precedes or follows another vowel in the same syllable. For example: *seis /seis/ → [seis] ‘six’; agua /agua/ → [‘a.ɣa] ‘water’*. The mid vowels /e/, /o/ also have the corresponding allophones [ɛ], [ɔ], which occur in colloquial speech, as in *área /area/ → [’a.ɾə] ‘area’, poesía /poesia/ → [poe.’si.a] ‘poetry’* (Hualde 2005: 54-55). Note that diphthongization does not occur between vowels of the same height and roundness (Harris 1983: 17). Aside from the gliding
process, Spanish vowels tend to maintain their phonemic properties of height, roundness, frontness/backness.

2.2.2 Vowels in General American English

The vowel phonemes of GAE are as follows in Fig. 2.2. The back high and back mid vowels are rounded, while all other vowels are unrounded (Giegerich 1992; McMahon 2002).

**Fig. 2.2 English vowels**

<table>
<thead>
<tr>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
</tr>
<tr>
<td>/i/</td>
<td>/u/</td>
</tr>
<tr>
<td>/e/</td>
<td>/o/</td>
</tr>
<tr>
<td>/æ/</td>
<td>/ɑ/</td>
</tr>
<tr>
<td>Low</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.2.3 Diphthongs in General American English

GAE also has three phonemic diphthongs: /aɪ/ mice, /aʊ/ pout, /ɔɪ/ ([ɔɪ] in some dialects) voice. In addition to these phonemic diphthongs, many vowels are characteristically realized as diphthongs or long vowels, as in Fig. 2.3 (Giegerich 1992; McMahon 2002).

**Fig. 2.3 Diphthongal or long realization of English vowels**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>peer [pʰɪːr]</td>
<td>moose [muːs]</td>
</tr>
<tr>
<td>pace [pʰeɪs]</td>
<td>coat [kʰoʊt], [kʰoːt]</td>
</tr>
<tr>
<td></td>
<td>caught [kʰɔːt]</td>
</tr>
<tr>
<td></td>
<td>pot [pʰɑːt]</td>
</tr>
</tbody>
</table>
2.2.4 Comparison of Spanish and GAE Vowel Systems

The following is a comparison of the vowel systems of Spanish and English, as can be seen in Fig. 2.4.

**Fig. 2.4 Comparison of Spanish and English Vowel Systems**

<table>
<thead>
<tr>
<th></th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>/i/ → [i]</td>
<td>/i/ → [i:]</td>
</tr>
<tr>
<td></td>
<td>/u/ → [u]</td>
<td>/u/ → [u:]</td>
</tr>
<tr>
<td></td>
<td>(when unstressed and adjacent to another vowel)</td>
<td>/i/</td>
</tr>
<tr>
<td>Mid</td>
<td>/e/ → [ɛ]</td>
<td>/e/ → [æ]</td>
</tr>
<tr>
<td></td>
<td>/o/ → [ɔ]</td>
<td>/o/ → [o:]</td>
</tr>
<tr>
<td></td>
<td>(only in colloquial speech)</td>
<td>/e/</td>
</tr>
<tr>
<td>Low</td>
<td>/a/</td>
<td>/æ/</td>
</tr>
</tbody>
</table>

Although many of the same phonemes are found in Spanish and English, they have different realizations. The high vowels /i/, /u/ are short and maintain a monophthongal pronunciation in Spanish, while their GAE counterparts are longer and have a diphthongal quality. /i/, /o/ are absent in Spanish and are shorter, lower, and more central than /i/, /o/ (Hualde 2005: 126).

The mid vowels /e/, /o/ maintain their vocalic quality in Spanish. However, they can have higher or lower realizations depending on dialect and phonetic context (Navarro Tomás 1932: 52-53, 59-60). In GAE, /e/ and /o/ are frequently realized as the diphthongs [ei] and [ou], respectively. /e/ has an especially consistent diphthongal

---

2 Navarro Tomás (1932: 52, 53) states that /e/ in contact with /i/ and /x/ will sometimes have a realization that approximates English /æ/, for example terremoto /teremoto/ [tr.re.ˈmo.to] ‘earthquake’, ejemplo /exemplo/ [e.ˈxem.plo] ‘example’. Likewise, /o/ can have an open pronunciation in the same phonetic context, but it does not as closely approximate English /o/ (59,60). Stockwell & Bowen (1965: 119) also claim that /e/ may approximate /æ/ in syllables with /s/ or /l/ in coda position, for example desde /desde/ [ˈdez.de] ‘from, since’, vuelto /buelto/ [ˈbuel.to] ‘return, returned – past participle’. The realization of the mid vowels in Spanish is also dialectal. For example, in Mexican Spanish, there is a tendency to raise mid vowels in hiatus to form diphthongs: teatro /teatro/ → [ˈtja.tro] (Barrutia & Schwegler 1994: 238). In Costa Rican Spanish, /e/ and /o/ have a tendency to raise to /i/ and /u/, respectively, and in Andean Spanish /e/ and /o/ often merge with /i/ and /u/ because of contact with indigenous languages (Lipski 1994).
pronunciation in GAE. The vowels /e/, /ə/ do not occur as phonemes in Spanish. As phonemes in GAE, they are lower and shorter than /e/, /o/. The mid central vowels [ə], /ʌ/ are absent from the Spanish vowel inventory (Hualde 2005: 125-126).

The low vowel inventories of the two languages do not share common vowels. Spanish has /a/ as its single low vowel. GAE has the low vowels /æ/ and /ɑ/.

While both languages have diphthongs as part of their vowel systems, they differ in their phonetic or phonemic status. Diphthongs in Spanish result from the gliding of the high vowels /i/, /u/ and are not phonemic. In English, there are both phonemic and phonetic diphthongs. The three phonemic diphthongs are /aɪ/, /aʊ/, /ɔɪ/, and the phonetic diphthongs are the frequent realizations of the mid vowels /e/ and /o/.

Another difference between Spanish and English concerns the relationship of stress and vowel pronunciation. Although this difference goes beyond the segmental level, it is worth mentioning here because it presents an additional allophone of the English vowel inventory that is not present in Spanish. This allophone is [ə], or schwa, and is the most common vowel in English. It occurs in unstressed syllables and can appear in any syllable within the word. It can also be the only vowel, or even the only segment, in an unstressed word, such as the indefinite article a. Because schwa only occurs in unstressed syllables and is in complementary distribution with the other vowels (no minimal pairs), it is not considered a phoneme in the GAE vowel inventory (Giegerich 1992: 68-69). Rather, it is an allophone of the other vowels. In contrast, Spanish vowels show little variation in their quality, whether stressed or unstressed (Hualde 2005: 126-127).
In summary, the vowels in Spanish’s smaller and simpler inventory are shorter and more faithful to their vocalic qualities than those found in GAE. GAE vowels tend to lengthen and form diphthongs. GAE has both phonetic and phonemic diphthongs, whereas Spanish diphthongs result from the gliding of the high vowels /i/, /u/. The pronunciation of vowels in GAE relies on whether they are part of a stressed syllable, whereas in Spanish, vowels maintain their phonemic qualities independent of stress. A more specific discussion of vowels and glides and their interaction within syllables is the topic of the next section. From this point, GAE will be referred to as English.

2.3 The Syllable in Spanish and English

This section will describe and compare syllable structure in Spanish and English. Allowable segments in each syllable constituent will be discussed.

2.3.1 The Syllable in Spanish

The syllable is an essential unit for understanding phonological processes in Spanish. Knowledge of syllable constituents and boundaries constitutes part of a speaker’s linguistic competence in their native language, and agreement on allowable constituents and boundaries must exist if the syllable is to be used as an organizational unit for understanding the distribution of phonemes. In Spanish, agreement on what is an allowable syllable constituent and on where the syllable boundaries fall is clear in comparison with English (Harris 1983: 4; Hualde 2005: 70). This section describes what are agreed upon as syllable constituents and boundaries in Spanish, especially concerning the behavior of vowels and glides.

The syllable has two main constituents, the onset (O) and the rhyme (R). The rhyme consists of two parts, the nucleus (N) and the coda (C). The nucleus is the only
obligatory part of the syllable, and the onset and the coda are optional constituents in Spanish (Hualde 2005: 71). Spanish syllables can contain maximally five segments, three of which would pertain to the rhyme, as its maximum number of constituents, as in Fig. 2.5 (Harris 1983: 9-10). The following sections will give the acceptable segments for each syllable constituent for the purpose of comparison with English syllables and to later anticipate areas of difficulty for the acquisition of Spanish syllables.

**Fig. 2.5 The Spanish Syllable**

```
       σ
      / \
     /   \
   (O)   R
     \   /
      \ / \
     N   (C)
```

*first syllable of transporte ‘transport’*

**Nucleus**

There are restrictions as to what is an acceptable nucleus in Spanish. The nucleus contains the most sonorous part of the syllable, and is therefore its most perceptible element. It is also the most open segment of the syllable, with the most intensity, and has the potential to be lengthened (Quilis 1993: 362). In Spanish, the nucleus is always a vowel. In fact, a single vowel can make up a syllable. For example, the word o ‘or’ consists of a single vowel, as does the first syllable in [a:mor] amor ‘love’. A glide may also occur in the nucleus, preceding the vowel, so that single vowels and rising diphthongs may occur as a syllable nucleus in Spanish (Hualde 2005: 70-71). Possible Spanish nuclei are summarized in Fig. 2.6.
Fig. 2.6 Possible Spanish Nuclei

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Obligatory constituent; single vowel is the most common constituent</th>
<th>[aˈ.mor] ‘love’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rising Diphthong</td>
<td>Glides may precede a single vowel</td>
<td>[seɪs] ‘six’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[ˈa.ɣua] ‘water’</td>
</tr>
</tbody>
</table>

Onset

There are a number of restrictions on allowable onset constituents. Onsets in Spanish allow for single consonants and for clusters of no more than two consonants. Any single consonant is eligible as a simple onset; however, there are restrictions governing allowable onset clusters (Harris 1983:14). Complex onsets may contain a plosive or /l/ followed by /l/ or /ɾl/. The only onset clusters that are allowed are ones that can also occur word initially. The last restriction concerns /s/. Since words cannot begin with /s/ + consonant in Spanish, /s/ + consonant is not an allowable onset cluster (Hualde 2005: 73-74). Fig. 2.7 summarizes possible Spanish onsets.

Fig. 2.7 Possible Spanish Onsets

<table>
<thead>
<tr>
<th>Simple onsets</th>
<th>Any single consonant</th>
<th>[ˈga.to] ‘cat’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex onsets</td>
<td>Maximally 2 consonants – plosive or /l/ + /l/ or /ɾl/ + C</td>
<td>[ˈbra.so] ‘arm’</td>
</tr>
<tr>
<td></td>
<td>*/s/ + C */dl/</td>
<td>[flor] ‘flower’</td>
</tr>
</tbody>
</table>

Coda

There are also restrictions on codas. Like onsets, codas may contain a single consonant or a cluster of no more than two consonants. In a cluster, the second consonant is obligatorily /s/. Coda consonants are more restricted than those allowed in the onset. The most frequent coda consonants are /dl/, /ls/, /n/, /l/, and /ɾl/, and all are subject to neutralization processes (Hualde 2005: 75-6). These neutralization processes are detailed in Fig. 2.8. With the exception of nasal assimilation, which occurs
independently of dialect and register, the reduction processes described above occur most often in rapid or casual speech, and some are dialect specific. Because of the existence of so many reduction processes, it is clear that coda consonants are in a weaker position than those in the onset. Furthermore, these reduction processes can be considered attempts at eliminating the coda in order to achieve the common CV syllable that is favored in Spanish.

In addition to consonants, post-vocalic glides are parsed in the coda and can be followed by /n/ or /s/ (Hualde 1992: 483). For example, the first syllable of [treɪn.ta] treinta ‘thirty’ contains a glide + /n/ sequence, and the last syllable of [a.βe.ɾi.ɣuaiʃ] averiguáis ‘you-pl find out’ has a glide + /s/ sequence. It may be worth pointing to the limited contexts of the rhymes found in the previous two examples. Glide + /n/ sequences only occur word-internally. A triphthong + /s/ sequence is limited to the vosotros forms of Spain and to the historically related voseo forms of some Latin American dialects (Harris 1983: 16; Lipski 1994). With such limited contexts for complex rhymes, it is no surprise that the simple CV syllable is the most commonly

---

3 Coda /s/ is aspirated or deleted in the Spanish of the following countries to varying degrees, depending on social class and register: Argentina, Chile, Cuba, Dominican Republic, coastal regions of Ecuador, Nicaragua, Panama, Paraguay, Puerto Rico, southern Spain, Uruguay, Venezuela. In some regions, aspiration or deletion of /-s/ is nearly categorical. It is possible to find /-s/ reduction in other regions not listed above but to a lesser degree and in limited contexts. Neutralization of /l/ and /ɾ/ can be found in Cuba, Dominican Republic, coastal regions of Ecuador, Panama, Puerto Rico, and Venezuela. Neutralization depends on social class, and the resulting sound depends on region. For example, in the Dominican Republic, the resulting sound is the glide [j] (Lipski 1994).

4 The use of vosotros (2nd person plural – informal) is limited to Peninsular Spanish. Voseo (in place of or alongside the use of tú) is in use throughout much of Latin America, and certain regions in the following countries present these verb forms with diphthongized endings: Bolivia, Colombia, Costa Rica, Ecuador, Panama, Venezuela (Lipski 1994).
occurring syllable in Spanish, making the most common rhyme that which consists of a single nuclear vowel with a null coda (Quilis 1993: 370).

**Fig. 2.8 Spanish Coda Neutralization and Reduction**

| Fricatives | /s/: aspirated or deleted – dialectal /ʃ/ x/: rare, aspirated or deleted | [ˈes.to] ~ [ˈeh.to] ~ [ˈe.to] ‘this’ [re.’lox] ~ [re.’lo] ~ [re.’lo] ‘clock, watch’ |
| Nasals | Assimilate to the following consonant | [ˈteŋ.go] ‘I have’ [in.] ‘influence’ |
| Liquids | /ɾ/ : neutralized – dialectal /ɾ/ : no contrast between trill and flap | [ˈʃuer.to] ~ [ˈʃuel.to] ‘port’ [a.‘blar] ~ [a.‘blar] ‘to speak’ |
| Glides | Post-vocalic: can be followed by /n/ or /s/ Glide + nasal consonant in the rhyme only occurs word internally | [seis] [ˈtreɪ.in.ta] |

### 2.3.2 The Syllable in English

As in Spanish, the syllable is crucial for understanding many phonological processes in English. Knowledge of syllable structure in English provides for a more accurate understanding of phoneme distribution, stress placement, and rhythm and intonation patterns. Being able to count the number of syllables in a given word is part of the linguistic knowledge of native speakers; however, many speakers of English would have difficulty pointing to the divisions between syllables (McMahon 2002: 104-105).

This section describes syllable structure in English by detailing permissible syllable constituents and generally agreed upon syllable boundaries.

The nucleus is the only obligatory syllable constituent in English (McMahon 2002: 105). While the onset and coda are optional constituents, it is possible and common to find comparatively heavy syllables in English with onsets and codas.
containing the maximum number of permissible segments (Shockey 2003: 32). Fig. 2.9 demonstrates an English syllable.

**Fig. 2.9 The English Syllable**

```
    σ
   /\  \\
  R  N  (C)
 (O) stɹ  ai  pt     striped
```

**Nucleus**

In English, the nucleus can be a vowel, a diphthong, or a consonant. As the nucleus is the only obligatory constituent, a syllable or a word may consist of a single vowel, such as the indefinite article \(a\) [ə]. In fact, [ə], as the most commonly occurring vowel in English, is the most common syllable nucleus. As was outlined in section 2.2.2, pronunciation of a single vowel in the nucleus is dependent on stress, contributing to the high occurrence of [ə]. This interaction between stress and vowel pronunciation will be further discussed in light of syllable weight in section 2.5.2. Concerning diphthongs, falling diphthongs are the only ones to appear in the nucleus, as English lacks rising diphthongs. The consonants that can be nuclear are the sonorants /l/, /m/, /n/, /r/.

For example, the words *little* and *cotton* have as their second syllables the nuclear consonants [l] and [n], respectively (McMahon 2002: 105). The allowable constituents are detailed below in Fig. 2.10.
Fig. 2.10 Possible English Nuclei

| Vowel | Any vowel permissible  
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[ə] most common</td>
</tr>
<tr>
<td></td>
<td>Pronunciation dependent on stress (see section ??)</td>
</tr>
</tbody>
</table>
|       | [hɪʔ] hit  [hæʔ] hat  
|       | [ə.ˈnæ.ɹ.ə.mi] anatomy |
| Diphthong | Only falling diphthongs |
|       | [mɑːs] mice, [pʰæʔ] pout  
|       | [pʰəς] pace |
| Consonant | Sonorants [l m n r] |
|       | [lɪd] little  [kɑʔn] cotton |

Onset

The onset may contain one, two, or three consonants, but there are restrictions on the possible consonant sequences that can fill this position. In a consonant cluster made up of three consonants, the first consonant must be /s/. /ŋ/ is not permissible in the onset. /v ð z ʒ/ do not appear in onset clusters. In addition, /t d θ/ /l/ are not permissible onset clusters (McMahon 2002: 105-106).

Although some onset reduction processes are pointed out in the figure below, the onset is fortified or created in some processes that concern vowels and glides. One change to the onset that occurs in casual speech is palatalization, which occurs when an alveolar stop or fricative is followed by /j/. For example, the phrase not yet would be pronounced as [næʧɛʔ] or would you as [wʊʤə] (Shockey 2003: 44-45). The onset is actually strengthened in this process. English also demonstrates a tendency to create onsets. Glottal stops are often inserted between two vowels or before a word-initial vowel that follows a consonant. Although glottal stop insertion is difficult to predict and varies by speaker and even by utterance, it demonstrates the motivation to have an onset in English, as do the examples in the figure below in which reduction tendencies do not
result in onset deletion (Olive, Greenwood & Coleman 1993: 128, 131). Possible onsets and related processes are summarized in Fig. 2.11.

**Fig. 2.11 Possible English Onsets and Related Processes**

<table>
<thead>
<tr>
<th>Simple Onset</th>
<th>Possible Onsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any consonant except /ŋ/</td>
<td>/ˈfeə.ðæ/ feather [əʊ.pi?] repeat</td>
</tr>
<tr>
<td>Glottal stop insertion before a vowel (varies by speaker and utterance)</td>
<td>[ʔeɪp] ape</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Complex Onset</th>
<th>Possible Onsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximally three consonants</td>
<td>/tɹ ɪp/ strip</td>
</tr>
<tr>
<td>In three consonant cluster: /s/ + CC</td>
<td></td>
</tr>
<tr>
<td><em>/v ð z ʒ/</em> + C</td>
<td></td>
</tr>
<tr>
<td><em>/t d θ/</em> + /l/</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Related processes</th>
<th>Possible Onsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>/h/ dropping</td>
<td>/dəzi/ does he</td>
</tr>
<tr>
<td>Palatalization: alveolar stop or fricative + /j/</td>
<td>[nɑʧ ɛʔ] not yet</td>
</tr>
</tbody>
</table>

**Coda**

The coda may also contain up to three consonants, not including the addition of a plural or possessive morpheme. Consonant sequencing in this position is subject to restrictions and reduction tendencies. /h/ is not permissible in coda position. In nasal + stop coda clusters, the constituents must have the same place of articulation (McMahon 2002: 105-106). Coda consonants experience various reduction tendencies. Word-final [v] is often deleted in spontaneous speech. The word of [ɔv] becomes [ə], and have [hæv], already reduced to [ɔv] in the contraction should’ve, is also reduced to [ə], for example shoulda in place of should’ve. In these examples, of and ‘ve are unstressed function words known as “weak forms” that are susceptible to reduction (Shockey 2003: 35, 46). While not necessarily a reduction process, /l/ is pronounced as [l], known as dark or velarized /l/ in coda position. This pronunciation is also found when /l/ is syllabic, so it can be said that dark [l] occurs in the rhyme (McMahon 2002: 110). When [l] occurs after a vowel, the acoustic difference between the vowel and the lateral consonant is minimized, creating a perceived codaless syllable (Olive, Greenwood &
Cohen 1993: 208). Restrictions on coda consonants and various reduction processes are
detailed in Fig. 2.12.

**Fig. 2.12 Coda Consonants and Reduction Processes**

<table>
<thead>
<tr>
<th>Restrictions</th>
<th>Nasal + stop (segments have same place of articulation)</th>
<th>[pʰʌmp] pump [ænt] rant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduction tendencies</td>
<td>*/h/ deleted in weak forms in spontaneous speech VNC cluster – vowel nasalized, nasal unarticulated /t/ glottalized before consonant or pause and after a nasalized vowel /d/ omitted between consonants, after /n/ /l/ velarization</td>
<td>[hæv] →[əv]→[ə] Should’ve → shoulda [hɑʔ] hunt [sæŋkæs] sandcastle [dʌl] dull</td>
</tr>
</tbody>
</table>

2.3.3 Comparison of Spanish and English Syllables

Spanish and English allow syllables with onsets, nuclei, and codas. They also allow syllables that consist solely of a nucleus with no onset or coda. The two languages also allow for onsetless syllables and, conversely, for codaless ones. The difference between the two languages is the composition of each of the syllable constituents. In general, English allows for heavier syllables than in Spanish with more permissible segments in all parts of the syllable. A comparison of syllable constituents in Spanish and English can be seen in Fig. 2.13.

In the nucleus, Spanish and English limit the type of segment allowed in this position. Both languages allow single vowels and diphthongs. The major difference between Spanish and English nuclei is that English has syllabic consonants, the sonorants [l m n r], as syllable nuclei, while Spanish allows only vowels or a vowel preceded by a glide. Both languages have a single vowel as their most common nucleus.
In onset position, both languages place restrictions on allowable constituents, especially concerning onset clusters. For single consonant onsets, Spanish has no restrictions. Any consonant in the phonemic inventory is eligible. For sole onset constituents in English, the only restriction is the prohibition of /ŋ/, which may only appear in coda position. Spanish and English onset clusters differ in the maximum number of permissible segments. Spanish allows up to two consonants in the onset, while English allows for up to three consonants.

As in the onset, English allows more consonants in the coda than in Spanish. English codas can have a maximum of three coda consonants, whereas Spanish codas may only have two consonants. Both languages place restrictions on permissible coda segments. The only segment that is barred from coda position in English is /h/. In comparison, Spanish has relatively few coda consonants, and they are all subject to reduction tendencies.

In summary, English onsets and codas allow more consonants than in Spanish, creating heavier syllables. Even in nuclear position, English allows consonants, where Spanish only permits vowels and vowel-glide combinations. Spanish allows few coda consonants and tends toward eliminating the coda, while English codas are not reduced to the same extent. This section has outlined rules and restrictions on syllable constituents. The next section will consider sonority, which provides for a more universal explanation of syllable structure.
Fig. 2.13 Comparison of Spanish and English Syllable Constituents

<table>
<thead>
<tr>
<th></th>
<th>Onset</th>
<th>Nucleus</th>
<th>Coda</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish</td>
<td>maximally 2 consonants – plosive or /l/ + /ɾ/</td>
<td>single vowel -- most common rising diphthong</td>
<td>Maximally 2 consonants, 2nd consonant always /s/ Most common: /d/, /l/, /ɾ/, /n/, and /l/ Post-vocalic glide Many consonants neutralized or deleted</td>
</tr>
<tr>
<td></td>
<td>*s/ + C</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>*dl/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.4 Sonority

Syllabification and the rules for allowable syllable constituents set forth in the previous sections can be further understood through the principle of sonority sequencing. Sonority can be thought of in terms of how much carrying power, or output, a sound has in relation to how much articulatory power goes into producing it. That is, more sonorous sounds, such as vowels, have high output produced with a low level of articulatory effort, and less sonorous sounds, such as stops, have low output produced with a high level articulatory effort. Sounds can be ranked according to this relationship (Clark & Yallop 1995: 61).

Sonority’s explanatory value for syllabification is based on the principle of the Sonority Cycle, which holds that the preferred syllable type is one that has a maximal rise at the beginning and a minimal or gradual drop at the end. In order to maintain the sonority cycle, the order of segments must adhere to the Sonority Sequencing Principle, which dictates the acceptable order of segments based on sonority for onset, nucleus, or
The segments are ranked so that those with the highest sonority ranking are placed closer to the nucleus, and segments with a lower sonority ranking are placed at the syllable margins. This principle is not a rule; rather, it expresses a cross-linguistic tendency. Indeed, onset and coda clusters that adhere to the Sonority Sequencing Principle are most common cross-linguistically, as will be discussed in section 2.4.2 (Clements 1990: 284-285).

Fig. 2.14 Sonority Scale

\[
O < N < L < V \quad (O = \text{obstruent, } N = \text{nasal, } L = \text{liquid, } V = \text{vocoid (glides and vowels)})
\]

A single sonority scale, such as appears in Fig. 2.14, represents sonority in all languages and is a part of universal grammar (Clements 1990: 294, 291). However, languages may differ from this sonority scale by allowing marked syllables that do not conform to the Sonority Sequencing Principle. Although marked syllables do make their way into individual languages’ syllable inventories, there does not seem to be a need for language specific sonority scales. As will be discussed in the following sections, the relative markedness of segments may account for any language specific constraints. Furthermore, the existence of language specific sonority scales weakens the universal explanatory power of the principle of sonority (Clements 1990: 294-296).

In spite of Clements’ (1990) view that there is no need for language specific sonority scales, further subdivisions do appear on sonority scales in the literature on individual languages. For Spanish, Hualde (2005: 72) suggests the following sonority scale, which appears in Fig. 2.15.
The literature for English phonology has also produced a more finely-grained sonority scale, such as the one in Fig. 2.16 based on Giegerich (1992) and McMahon (2002).

The subdivisions on the sonority scales presented above are useful for making more precise observations on the behavior of sounds within the syllable and, in turn, more precise cross-linguistic comparisons. The following section will discuss how sonority affects the nucleus in English and Spanish, as the nucleus is the general domain for diphthongization.

2.4.1 Nucleus

This section will describe the nucleus, which is the sonority peak of the syllable, in both Spanish and English. The behavior of glides and subdivisions among vowels on the sonority scale will be considered.

**Spanish Nucleus**

In Spanish, vowels, which are the most sonorous segments, make up the nucleus along with the glides that may accompany the vowel. In other words, segments in
sonority levels 6-4 on the proposed sonority scale in Fig. 2.15 are the permissible members of syllable peaks in Spanish (Hualde 2005: 72).

The subdivisions along the sonority scale according to vowel height are useful for explaining vowel behavior in Spanish. The lower sonority ranking for mid and high vowels in relation to low vowels indicates their susceptibility to gliding, as discussed in section 2.2.1. While all vowels are permissible as single elements of the nucleus, high and mid vowels are also susceptible to gliding when in contact with other vowels. As glides, these vowels are relegated to satellite status in the nucleus and cannot be the absolute peak of the syllable. This subdivision also explains high and mid vowels’ respective probability to be relegated to satellite status. Mid vowels can optionally become glides in casual speech, which means that they can maintain their status as the sole segment in the nucleus when in contact with another vowel, maintaining a hiatus. However, high vowels glide as a rule when in contact with another vowel. The higher sonority ranking for mid vowels points to their increased ability to maintain absolute peak position over high vowels when adjacent to another vowel.

The fact that high vowels and glides share the same sonority ranking points to their similar underlying phonemic value, the most common glides [i u] being allophones of high vowels. Furthermore, the fact that glides reside at the lowest sonority rank for allowable nuclei reflects their status as marginal elements within the nucleus, that when they do occur can never participate as the sole segment in the nucleus.

**English Nucleus**

In English, sonority is not the only explanatory principle for syllabification in the nucleus, as syllable weight and stress placement play a role, as well. However, sonority
is at play in the relationship between glides and high vowels and explains the existence of syllabic consonants in English.

The subdivision between high and low vowels on the sonority scale in Fig. 2.16 is illustrative of the relationship between high vowels and glides in English. Glides and high vowels are not phonetically distinct; rather, it is how they function in syllables that differentiates them. Vowels can be syllable peaks, whereas glides are found only at the syllable margins. For example, the word ye has [j] in the onset and [i:] in the nucleus. Similarly, the word woo has [w] in the onset and [u:] in the nucleus (Giegerich 1992: 94; McMahon 2002: 106). The same holds for words like you and we, in which the glide is in onset position and the vowel in the nucleus. The fact that glides and high vowels are close on the sonority scale reflects their same underlying phonemic value (discussed in section 2.6.1), but the fact that they hold different ranks on the sonority scale illustrates their differing status in relation to the nucleus.

With the exception of glides, liquids and nasals are the sounds closest to vowels on the sonority scale. In English, liquids and nasals can be [+syllabic], which allows them to be in nuclear position (McMahon 2002: 107). In principle, any segment can belong to the nucleus, as some languages even allow obstruents in this position, but this ability is determined by its rank on the sonority scale (Clements 1990: 293-294). In English, liquids and then nasals are the least sonorous sounds allowed in the nucleus.

2.5 Syllabic Processes

This section will expand on what has been described in the previous sections concerning syllable constituents by examining syllable processes in Spanish and English. Some of the processes discussed in this section explain syllabification of individual
words and offer explanations for the permissibility of certain segments over others, especially concerning the relationship of stress to allowable segments in the nucleus in English. Other processes go beyond the syllabification of single words and examine what happens across word boundaries. This section will begin by describing syllable processes in Spanish and English and will end by summarizing and comparing the processes in both languages.

2.5.1 Syllabic Processes in Spanish

CV Rule

The CV Rule is a basic rule of Spanish syllabification, and it is motivated by the tendency in Spanish to favor CV syllables. The CV rule is without exception in Spanish and states that a consonant is obligatorily syllabified with a following vowel. Hence, a word like hago /ago/ ‘I do/make’ will be syllabified as [ˈaɣo] to maximize the number of CV syllables (Hualde 1989: 821; 2005: 73).

Resyllabification

Resyllabification is an extension of the CV Rule that occurs across word and prefix boundaries. Under this process, a consonant at the end of a word is resyllabified with a vowel at the beginning of the following word, so that syllable boundaries and word boundaries do not always coincide. This process happens in connected, casual speech (Harris 1983: 43; Hualde 1989: 821-822; 2005: 87). For example, los autobuses ‘the buses’ would be syllabified as [lo.sau.to.βu.ses]. The -/s/ from los is resyllabified as the onset of the initial syllable of the following word. Coda consonants may still experience weakening, depending on the dialect, but weakening is not as common preceding a word-initial vowel (Hualde 2005:88).
Resyllabification only occurs with C-V sequences. C-C sequences are not affected by the resyllabification rule, and neither are word-initial glides. Word-initial glides are considered consonants in the context of this rule, and experience various levels of fortition depending on the dialect. For example, *el hielo* /el ielo/ $\rightarrow$ [el.ʝe.lo] ‘ice’; *con huevos* /kon uebos/ $\rightarrow$ [kon.ɡue.βos] ‘with eggs’ (Hualde 2005: 89). Through glide fortition, the glides become fricatives and function as consonants, thus creating an onset. Because C-C sequences are barred from resyllabification, glide fortition must occur before the process of resyllabification (Hualde 1989: 825-826).

**Contraction**

Syllable contraction is another form of resyllabification that occurs across word boundaries in connected speech and that affects vowels. It most commonly happens when an unstressed vowel at the end of a word is followed by another unstressed vowel at the beginning of a following word. For example, *la historia* /la istoɾia/ $\rightarrow$ [lais.ˈto.rja] ‘the story/history’; *tu amigo* /tu amigo/ $\rightarrow$ [tua.ˈmi.yo] ‘your friend’. In this process, the high vowel becomes a glide. The two words contract by forming a diphthong. Mid vowels are also affected by this process, becoming glides, for example *se hace* /se ase/ $\rightarrow$[ˈse a.ɾe] ‘one does’ or ‘it is done – passive’. Going a step further, mid vowels may heighten and become high glides in colloquial speech, as in [ˈsɾa.se] $\rightarrow$[ˈɾja.se]. If there are two vowels of the same height (high or mid), the first vowel will form a glide (Hualde 2005: 89-90).

Just as two vowels can contract to form a diphthong, it is also possible for three vocoids to contract into a single syllable by forming a triphthong. Contraction between three vocoids can only occur when the middle vowel is lower than the two vocoids.
surrounding it. For example, the phrase *tarea importante* ‘important work’, pronounced [ta.ˈɾe.a.im.por.ˈtaŋ.te] in careful speech, will contract to [ta.ˈɾeim.por.ˈtaŋ.te] in casual speech. Likewise, the phrase *bibliotecario italiano* ‘Italian librarian’, pronounced [bi.bljo.te.ˈka.rjo.i.ta.ˈlja.no] in careful speech, will be pronounced [bi.bljo.te.ˈka.ɾjo.ta.ˈlja.no] in casual speech. It is acceptable for these sequences to contract into a single syllable because the syllable nucleus, the lowest vowel, remains the most sonorous element of the syllable. In contrast, sequences in which the middle vocoid is the highest and least sonorous vocoid in the sequence do not contract into a single syllable. Because a single syllable cannot have two sonority peaks, a sequence in which two vocoids would form sonority peaks around a less sonorous vowel is not permissible. For example, the three vowel sequence in the phrase *habla y escribe* /ablə i eskribe/ will be reduced to two syllables instead of a single one, [ˈa.βla.yes.ˈkri.βe] ‘speak and write – 3rd person present’ (Navarro-Tomás 1932: 150-151; Quilis & Fernández 1997: 150; Hualde 2005: 93-94). Sonority contours created by contraction can be seen in Fig. 2.17 below.

**Fig. 2.17 Sonority contours resulting from syllable contraction**

![Sonority contours](image)

Elision of unstressed /e/ is another way in which syllables contract. Unstressed /e/ can be elided when in contact with another vowel. Elision occurs when /e/ is word-final, as in *est(e) hombre* ‘this man’, and when it begins a closed syllable, as in *ya (e)stamos* ‘we are already here’ (Hualde 2005: 90).
When two identical vowels appear in a row, they are often reduced to a single vowel. It is possible that up to three identical unstressed vowels in a row can reduce to the duration of a single vowel. When two vowels appear in a row, it is not necessary for them to be unstressed in order to contract. Contraction is especially likely to happen when the words have a close syntactic relationship, such as between object pronouns and verbs (Hualde 2005: 90-91). Some examples of contraction between identical vowels are: *está arriba* [es.ta.ɾi.ba] → [es.ta.ɾi ба] ‘is upstairs’; *lo ordeno* [lo or.ɾe.no] → [loɾ.ɾe.no] ‘I order it’. The vowel processes listed above, both glide formation to resolve word-medial hiatus and syllable contraction through vowel deletion, work to create a more preferred syllable type that conforms to the preferred sonority contour (Clements 1990: 301). Syllabic processes in Spanish are summarized in Fig. 2.18.

**Fig. 2.18 Syllabic Processes in Spanish**

<table>
<thead>
<tr>
<th>CV Rule</th>
<th>hago /ago/ → [a.ɣo]</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Consonant syllabified with a following vowel</td>
<td></td>
</tr>
<tr>
<td>-Maximize number of CV syllables</td>
<td></td>
</tr>
<tr>
<td>Resyllabification</td>
<td></td>
</tr>
<tr>
<td>-A consonant at the end of a word is resyllabified with a vowel at the beginning of the following word</td>
<td>los autobuses → [lo.sau.to.ɾu.ses] el hielo /el ielo/ → [el.ɾe.lo] con huevos /kon uebos/ → [koŋ.ɾe.ɾu.ɾu]</td>
</tr>
<tr>
<td>-Glides are considered consonants in the context of this rule and experience fortition</td>
<td></td>
</tr>
<tr>
<td>Contraction</td>
<td>la historia /la historia/ → [laɾi.to.ɾa] est(e) hombre bibliotecario italiano [bi.ɾi.to.ɾe.ɾi.to.i.ta.ɾi.no] → [bi.ɾi.to.ɾe.ɾi.to.i.ta.ɾi.no]</td>
</tr>
<tr>
<td>-Syllables contract via: diphthongization, elision, triphthongization</td>
<td></td>
</tr>
</tbody>
</table>
2.5.2 Syllabic Processes in English

Onset Maximalism

Onset Maximalism is a principle that governs syllable division. It states that as many consonants as possible should be assigned to the onset and as few as possible should be assigned to the coda. While adhering to this principle, it is necessary to maintain well-formed syllables. For example, in the word reader, the [d] would be parsed as the onset of the second syllable. Both syllables, [ɹi:] and [dǝɹ], are acceptable syllables. Also, in the word cloister, the group [st] would be parsed as the onset of the second syllable, as [st] is an acceptable onset in English in words such as stripe and strut. However, in a word like flounder, the group [nd] cannot be parsed in the onset, as [nd] is not an acceptable onset cluster. Such a coda cluster would be permissible according to the Sonority Sequencing Principle; however, onset maximalism will force the [d] to onset position, so that the word is parsed as such, floun.der. Onset Maximalism also holds across word and prefix boundaries. For example, in the phrase red alert, the [d] would shift to form the onset of the second word in the phrase (McMahon 2002: 111-112).

Onset Maximalism is not limited to C-V sequences. C-C sequences may combine to form onset clusters, such as [tɔ.k اللا.دائي] for talk louder in which the /k/ is parsed as part of the onset of the following word. However, this type of linking is not always attested and is dependent on style and tempo of speech (Giegerich 1992: 280-281).

Onset Maximalism can be further explained in terms of sonority and a preference to maintain the sonority cycle. In the word flounder, the [d] represents a sonority trough. Hence the lowest sonority point is chosen as the onset of the second syllable of the word, floun.der. Also in the example red alert, the [d] is the lowest point in sonority, and is the
onset of the second syllable, \textit{re.da.lert}. Starting the syllable at the lowest point in sonority allows for the onset length to be maximized and also allows for the maximal rise in the onset desired by the preferred sonority contour (Geigerich 1992: 169; Clements 1990: 300). /s/ presents an exception to the preferred sonority profile, as in the example \textit{cloister} in which the group [st] is parsed as the onset of the second syllable even though it violates sonority sequencing, \textit{cloi.ster}. Examples of these sonority contours are shown below in Fig. 2.19.

\textbf{Fig. 2.19 Sonority contours resulting from Onset Maximalism}

\begin{center}
\begin{tabular}{lll}
\textit{floun.der} & \textit{re.da.lert} & \textit{cloi.ster} \\
\end{tabular}
\end{center}

\textbf{Syllable Weight}

English has both light and heavy syllables. A syllable’s weight depends on the segments in the nucleus and the coda. A light syllable is a codaless syllable consisting only of a short vowel. For example, the first syllable of the word \textit{repeat}, [ræ], is a light syllable, containing only a short vowel with no coda. A heavy syllable is one with a complex rhyme. A complex rhyme can consist of a short vowel, a long vowel, or a diphthong in the nucleus with one or more consonants in the coda. Alternatively, a complex rhyme can have a branching nucleus, containing a long vowel or a diphthong, with no coda consonant. For example, \textit{bed} [bed], \textit{beast} [biːst], and \textit{bound} [baʊnd] all have a complex rhyme with a short vowel, a long vowel, or a diphthong, respectively, with one or more coda consonants. Words like \textit{pie} [paɪ], and \textit{pea} [piː], both with branching nuclei and no coda consonant, are heavy syllables, as well (McMahon 2002: 114).
Syllable weight is indicative of stress placement in English. Light syllables do not carry stress. Since all lexical words must carry stress, no lexical word in English contains only a short vowel in the nucleus. Hence, there are no such words as *[pɪ] or *[pɛ] in English. Light syllables do appear as unstressed syllables in polysyllabic words, as in the above example *repeat [ɹe piːt] (Geigerich 1992: 146; McMahon 2002: 114).

Syllabic processes in English are summarized in Fig. 2.20.

**Fig. 2.20 English Syllabic Processes**

<table>
<thead>
<tr>
<th>Onset Maximalism</th>
<th>-As many consonants as possible should be assigned to the onset and as few as possible should be assigned to the coda.</th>
<th>rea.der cloi.ster re.dal.ert floun.der</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syllable Weight</td>
<td>-Light syllables do not carry stress. - Since all lexical words must carry stress, no lexical word in English contains only a short vowel in the nucleus.</td>
<td>*[pɪ] *[pɛ]</td>
</tr>
</tbody>
</table>

**2.5.3 Comparison of Spanish and English Syllabic Processes**

This section will compare syllabic processes in Spanish and English. Syllabic processes in Spanish and English have similar motivations, in that both languages work towards maximizing the onset. The CV Rule and the process of resyllabification in Spanish create an onset by attaching a consonant to a following vowel at the word and phrase level. Contraction also works to maximize the onset through diphthongization across words and vowel elision in Spanish. The principle of Onset Maximalism in English seeks to have the maximum number of segments possible in the onset. However,
English Onset Maximalism is carried a step further than in Spanish. In Spanish, resyllabification only occurs between C-V sequences, so that what appears to be Onset Maximalism in the CV Rule does not extend to resyllabification (Hualde 1989: 829). Onset Maximalism in English is more extensive in that consonant clusters may be created across word boundaries. Both languages do not seek to maximize the onset if doing so will result in ill-formed syllables. Maximizing the onset is not necessarily a language specific rule; rather, its motivation may be to achieve the universally preferred sonority profile of the CV syllable (Clements 1990: 317).

The languages differ in the relationship between stress and syllable formation. Syllable weight and stress placement influence what is considered to be a well-formed syllable in English. In Spanish, syllable well-formedness is largely independent of stress. The next section will discuss syllable structure and syllabic processes in so far as they can be used to explain the phonemic status of glides in both Spanish and English.

2.6 The Phonemic Status of Glides

In both Spanish and English, the phonemic status of glides has been a point of debate for phonologists. The cause for debate is the ambiguous status of glides as something in between a consonant and a vowel. Glides behave like a consonant in some contexts and as a vowel in others. As part of the syllable onset, they are considered consonants, while as part of the syllable peak, they interact closely with vowels. Phonetically, glides are more like vowels, but they can also be strengthened to be more consonant-like. This section will discuss the status of glides in Spanish and English along with the behavior of glides in different parts of the syllable.
2.6.1 The Status of Glides in Spanish

This discussion of the phonemic status of glides in Spanish begins with an explanation of exceptional hiatus, which presents a challenge to the gliding rule by which glides are derived. The discussion then continues on to the possible phonemes and allophones associated with glides in Spanish.

Exceptional Hiatus in Spanish

Words with an exceptional hiatus pronunciation present a challenge to the gliding rule in Spanish. Recall from section 2.2.1 that the gliding rule produces [i u] when the high vowels /i u/ are unstressed and next to a vowel that does not share the same height and roundness. However, there are lexical exceptions to the gliding rule in which we find alternation between diphthongs and hiatus in the same phonetic context. That is, a word with an expected diphthongal pronunciation according to the gliding rule may also have an optional hiatus pronunciation. The choice of pronunciation between hiatus and diphthong varies with style and dialect. Words with an optional hiatus are less frequent than those that are consistently pronounced with a diphthong, so that most words with a diphthong do not have an optional hiatus pronunciation. The words that do have an optional hiatus pronunciation are exceptions to the gliding rule (Hualde 1997; 2005: 80-82). Fig. 2.21 demonstrates pairs of words that differ in having an exceptional hiatus pronunciation or a consistently diphthongal pronunciation, based on native speaker perception. All examples in this section are from Hualde (1997, 2005) unless otherwise noted.
Fig. 2.21  Exceptional hiatus/consistent diphthong pairs (Hualde 1997: 64; 2005: 81)

<table>
<thead>
<tr>
<th>Hiatus</th>
<th>vs.</th>
<th>Diphthong</th>
</tr>
</thead>
<tbody>
<tr>
<td>dueto [du.eto] ‘duet’</td>
<td></td>
<td>duelo [dŭelo] ‘duel’</td>
</tr>
<tr>
<td>cliente [kli.en.te] ‘client’</td>
<td></td>
<td>diente [dĕnte] ‘tooth’</td>
</tr>
<tr>
<td>pié [pi.e] ‘I chirped’</td>
<td></td>
<td>pie [pię] ‘foot’</td>
</tr>
<tr>
<td>riendo [ri.endo] ‘laughing’</td>
<td></td>
<td>siendo [sĕiendo] ‘being’</td>
</tr>
<tr>
<td>huimos [u.imos] ‘we flee’</td>
<td></td>
<td>fuimos [φuimos] ‘we left, we were’</td>
</tr>
<tr>
<td>enviamos [emvi.amos] ‘we send’</td>
<td></td>
<td>envidiamos [emviði.amos] ‘we envy’</td>
</tr>
</tbody>
</table>

While the hiatus pronunciation is exceptional, it is not unpredictable. There are contexts in which lexical hiatus is more likely to occur. A hiatus pronunciation is possible if a morphologically related word has the high vowel as its stressed vowel, for example env[i.a]mos ‘we send/sent’ from envía ‘send – 3rd person singular’ or [o.i]ré ‘I will hear’ from oír ‘to hear’ (Hualde 1997: 67; 2005: 82).

Another context in which the hiatus pronunciation is possible is if a morphological boundary exists between the two vocoids. A hiatus pronunciation is more likely with compounds, prefixes b[i.e]nio ‘two year period’, suffixes virt[u.o]so ‘virtuous’ (Hualde 1997: 67; 2005: 82-83).

There are also two conditions that predict the use of exceptional hiatus. The Initiality Condition provides for optional hiatus to be more probable at the beginning of the word, for example b[i.o]logo ‘biologist’ (Hualde 1997: 67; 2005: 84). The Stress Condition provides for optional hiatus if the second vowel in the sequence is stressed or if the syllable following the vocoid sequence is stressed, for example in the following forms of dialogue pronounced with an exceptional hiatus: d[i. ‘a]logo ‘dialogue’, d[i.a]. ‘lo.go ‘I dialogue’. However, dialogó d[i.a].lo. ‘go ‘dialogued – 3rd person singular’ is pronounced with a diphthong because the stressed syllable is removed from the vocoid
sequence (Hualde 2005:84). In addition, exceptional hiatus occurs most often when the 
vocoid sequence has increasing sonority and in /ia/, /io/ sequences (Hualde 1997: 67;  
2005: 84).

The conditions and opportunities for analogy listed above do not necessitate a 
hiatus pronunciation; they are contexts in which exceptional hiatus may occur (Hualde  
2005: 84) However, lexical hiatus is subject to dialect variation and can even vary 
between individuals. For example, lexical hiatus occurs more frequently in Castilian  
Spanish than in Latin America (Hualde 2005: 80-82). Style can also determine the  
likelihood of exceptional hiatus. Words with a high + non-high vowel sequence, in 
which the non-high vowel is stressed, have a hiatus pronunciation in careful speech, 
while a diphthongal pronunciation is demonstrated in rapid speech (Navarro-Tomás  
1932:155). Slow, emphatic speech can also lead to hiatus pronunciation, especially when 
the sequence is in word-final position, such as in santuario ‘sanctuary’, prior, embrión  
‘embryo’ (Navarro-Tomás 1932: 158-159).

The existence of an exceptional hiatus pronunciation throws the gliding into 
question. While there is no doubt that diphthongization through the gliding rule is the 
preferred way to resolve the more marked hiatus sequence, there is still a preference in 
some dialects and for some individuals to maintain the marked structure in certain  
contexts. The next section will address the phonemes and allophones involved in the 
behavior of glides in Spanish.

**Phonemes and Allophones of Spanish Glides**

Considering diphthongs allows for an examination of the phonemic status of 
glides in Spanish. In Spanish, diphthongs are composed of two vowel phonemes, and
triphthongs are composed of three vowel phonemes (Alarcos Llorach 2007: 160). For example, the underlying representation for gracias ‘thanks’ is /grasias/; likewise, the underlying representation for buey ‘ox’ is /buei/. The glides are derived via the gliding rule, as discussed in section 2.2.1, which has as its only exceptions the words with an optional lexical hiatus pronunciation, as discussed in the previous section. Hence, the status of the glides [i] and [u] as allophones of the phonemes /i/ and /u/, respectively, is relatively uncontroversial.

While the fact that glides are derived from vowels is generally uncontested, the allophones that are associated with the high vowels are up for debate. It has been proposed that the phoneme /i/ has the allophones [i], [ɾ], [j], while /u/ has the allophones [u], [u], [w]. The way that these allophones function in the syllable is what differentiates them. [i] and [u] are always syllable nuclei. [i], [u], [j] and [w] are found at the margins of the nucleus and are differentiated by their pre- or post- nuclear status. Under this proposal, [j] and [w] are the first segments of a rising diphthong and are considered semi-consonants, for example quiero [kjeɾo] ‘I want’, puedo [pweðo] ‘I can’. [i] and [u] are the last segments of a falling diphthong and are considered semi-vowels (Quilis 1993: 179-180; Quilis & Fernandez 1997: 70; Alarcos Llorach 2007: 152-153). Hualde (1997) makes no distinction between on-glides and off-glides. Instead, he proposes the single allophone /j/ for both positions, for example /diente/ [djente] ‘tooth’ and /peine/ [pejne] ‘comb’ (Hualde 1997: 68). Although his investigation concerns /i/, one can assume that the glide resulting from /u/ would also be assigned one allophone for both positions. It is tempting to assign semi-consonant status to a pre-nuclear glide because of their behavior in onsets, which will be discussed in the next paragraph. However, Hualde’s analysis is
better motivated in that both on- and off-glides are derived by the same phonological process resulting from their adjacency to another vowel.

The behavior of glides as syllable onsets in words such as *yema* ‘yolk’ and *callar* ‘to quiet’ provides another question. In these words, the glide functions as a syllable onset. If glides are allophones of vowels, then there is a conflict as vowels can only form part of the nucleus, not the syllable onset. One solution to this conflict is to propose the phoneme /y/ that has overlapping allophonic realizations with /i/. For example, *reyes* ‘kings’ and *mayo* ‘May’ have differing phonemic representations of the glide, /reies/ and /mayo/, respectively. However, both words are pronounced with the allophone [y], [reyes] and [mayo]. /y/ also has the affricate allophone [ŷ] that appears word-initially, as in the above example *yema* (Alarcos Llorach 2007: 163).

In another response to the onset question, Hualde (1997) proposes the phoneme /i/ with the allophones [ʝ] ([ʃ]), and [ʝ] ([ʃ]) in onset position. He argues that consonantization creates the voiced palatal continuant that appears when /i/ is in syllable initial position (Hualde 1997: 68-69). Syllable contraction also results in fortition of /i/, for example *el[ʝ]elo* from *el hielo* ‘ice’ (Hualde 1997: 66). Some evidence against the existence of a single phoneme /i/ is found in Argentinian Spanish, in which speakers make a phonemic distinction between the glide and the obstruent consonant. In some dialects of Argentinian Spanish, speakers use the allophone [ʒ] or [ʃ] only for orthographic *y* and *ll*. In other contexts the allophone [ʝ] is used (Hualde 1997: 72-73; 2005: 166). The phonemic difference between [ʝ] and [ʒ] is illustrated with the minimal pair *hierba* ‘mate tea’ and *yerba* ‘grass’. *Yerba* would be pronounced as [ʒ]erba, whereas
hierba would be pronounced as \[ \text{[i]erba} \]. (Hualde 2005: 169). The two distinct phonetic realizations reveal a phonemic distinction between /\text{ʒ}/ and /i/ (Hualde 1997: 73).

In yet another attempt to solve this onset conflict, Hualde proposes the “quasi-phoneme” /\text{j}/ (2005: 171). This symbol is used to represent a voiced palatal consonant with variable constriction, which ranges from a glide, to a stop, to an affricate. This sound is used when pronouncing syllable-initial y and ll, as represented orthographically (Hualde 2005: 165). One can observe a “quasi-contrast” between words like yema and hielo. Words with an initial hiV sequence are often not strengthened to the same extent as words with an initial orthographic y. While yema can be pronounced as [iema]~[jema]~[jema], hielo shows less constriction, [ielo]~[jelo] but not [jelo]. There is also a “quasi-contrast” found in huV words and those beginning with gu, for example hueso ‘bone’ vs. guasa ‘humor’. Fortition will create the pronunciation [gueso] that alternates with [gueso]. The fortified pronunciation is common except for after a pause when the contrast is maintained. These contrasts and the different levels of constriction are dependent upon dialect and speech style, so that the distinction between /i/ and /j/ is not complete. At this point, the phonemic status of /j/ is still up for debate (Hualde 2005: 169-172).

2.6.2 The Status of Glides in English

In English, glides and high vowels are not phonetically distinct; rather, it is how they function in syllables that differentiates them. The glides [j] and [w] are allophones of the high vowels /i/ and /u/, respectively, and they have differing values for the feature [syllabic]. Glides are [-syllabic], and the high vowels are [+syllabic]. Hence, glides and high vowels are in complementary distribution. Vowels can be syllable peaks, whereas
glides are found only at the syllable margins. For example, the word ye has [j] in the onset and [i:] in the nucleus. Similarly, the word woo has [w] in the onset and [u:] in the nucleus (Giegerich 1992: 94; McMahon 2002: 106). The same holds for words like you and we, in which the glide is in onset position and the vowel in the nucleus.

/j/ and /w/ are also proposed as phonemes in English, /j/ being the more interesting of the two glides. /j/ is permissible as a single onset, such as in yard. However, its behavior in onset clusters is more problematic. It occurs with nonsonorants more frequently than other sonorants. It can even occur with nonsonorants that normally do not appear in clusters, such as hue, pew, few, beautiful, cue, view (some examples are Giegerich’s). It can also occur after sonorants, for example music. As a single onset consonant, /j/ can occur before any vowel, but as part of a cluster, it can only appear before /u/. Because /j/ and /u/ have such a close relationship, they must form a phonological unit. Therefore, /j/ must be part of the nucleus in these examples. If /j/ is part of the nucleus, then it is subject to the rules governing onset constituents (Giegerich 1992: 157).

However, /j/ must also belong to the onset. If it were not part of the onset then other permissible onsets could occur before it, such as other permissible onset clusters like */strju/ or */klju/. It clearly occupies a spot in the onset in that it occurs only after a single consonant or after an sC onset cluster. Geigerich proposes that /j/ is part of the onset and the nucleus (1992: 158). The fact that /j/ cannot occur after coronal consonants also demonstrates its joint association with the onset and nucleus. This is a constraint that does not hold for other onset segments that are not also part of the nucleus (Geigerich 1992: 159). This proposal does not settle the debate about the syllabic affiliation of
glides in English. The phonemic status and the syllabic affiliation of glides in the
literature is still up for debate.

2.7 Making Predictions

This section will make predictions as to which aspects of Spanish diphthong
pronunciation will most likely cause difficulty for L2 learners of Spanish with English as
a native language. The predictions made in this section are based on the descriptions and
comparisons offered in this chapter.

2.7.1 Predicting Errors in SLA

Many of the predictions that have been made as to the relative difficulty of
acquiring L2 structures have been made according to the framework of Contrastive
Analysis (CA). Within this framework, it is held that one of the central obstacles to
acquiring an L2 is structural interference from the L1. The differences between the L1
and L2 are a major source of interference for the language learner. Hence, it is necessary
to carry out a contrastive analysis of the two languages to identify what will be the most
difficult aspects of acquisition on which the researcher, instructor, and learner should
focus. Carrying out a contrastive analysis of English and the five most commonly taught
foreign languages in the United States in the middle of the twentieth century was the

While it has been approximately half a century since Stockwell & Bowen (1965)
undertook their pioneering series, predictions based on L1 and L2 comparisons are still
useful for L2 researchers and instructors, as is evidenced by more recent investigations
that still rely on comparisons of L1 and L2 phones as the basis for making hypotheses for
L2 speech errors. Even in early theories of interlanguage that began to move away from
a contrastive approach, Corder (1967) points to the fact that a considerable amount of learners’ errors are related to L1 interference, and Selinker (1972) includes the L1 as an active system in L2 production. Moving forward, Eckman (1987) argues against the complete abandonment of a contrastive framework. He claims that the comparison of the L1 and L2 is “crucial” in predicting the degree of difficulty that a language learner might experience (Eckman 1987: 55). Other studies over the recent decades take phonemic comparisons as a point of departure for empirical work (Flege & Hillenbrand 1984; González-Bueno 1997; Díaz-Campos 2004; among many others). Archibald (1998, 2000) and Young-Scholten & Archibald (2000) while not studying particular phones or languages, claim L1 transfer to be a dominant process in L2 phonological acquisition. These studies are examples of the many that exist in the field of SLA based on L1 and L2 comparisons. However, structural differences between the L1 and L2 are a point of departure for these studies; the difference itself is not the explanatory principle as it is in the work of Stockwell & Bowen (1965).

Although L1 influence is able to explain some L2 pronunciation errors, it cannot explain all of them. It is evident that other explanations are necessary (Eckman 2004: 517). As early as Corder it was clear that contrastive analyses of L1 and L2 systems meant to aid the language instructor in identifying problem areas fell short and that many errors were not predicted by these analyses (1967: 162). Subsequently, there was a change to reliance on developmental processes that focused on the language learner’s hypotheses and learning strategies. Under this new development, errors were seen as proof that the learner is investigating the L2 system and as evidence of learning strategies (Corder 1967: 168). One of the weaknesses of CA was that it did not recognize the
learner’s cognitive skills. Under the behaviorist CA model, the learner was conditioned by the L1 and was not capable of creating forms that departed from it. However, it is evident that learners create new forms that are not accounted for in the CA model. In fact, learners produce forms that do not exist in the L1 or the L2, so that they are not solely conditioned by the two languages in question. Flege & Hillenbrand show that learners may actually produce new L2 segments better than those that have L1 counterparts (1984: 717). Hence, another weakness of CA is its inability to predict forms that are not part of the L1 or L2 inventory, nor is it able to predict the success with which learners may produce new L2 forms through establishing new phonetic categories.

An additional weakness of CA is that it is too simplistic. It is generally concerned with comparisons between L1 and L2 phones and does not consider higher order linguistic processes. Work by Archibald (1998) is influential in demonstrating the need to view the interaction between the L1 and L2 as going beyond the analysis of specific segments. Archibald (1998) favors a phonological theoretical framework that captures the mental representations of L2 learners. Archibald (1998) views L1 transfer as the dominant process in IL phonology; however, he does not refer to a superficial segmental transfer. Instead, the learner transfers complex mental representations and L1 principles and parameters. It is not only structural changes that are taking place; hierarchical representations are transferring (1998: 209, 192). Although the transfer of principles and parameters is the dominant process, Archibald’s later work recognizes the important role of universal markedness (2000: 127). L1 segments do transfer, but they are part of a larger representational scheme for which CA cannot account. It has also been
acknowledged through the work pertaining to L1 transfer that other models, such as Markedness, need to be incorporated to account for all L2 errors.

Eckman (1987) argues for Markedness to be incorporated into the CA model. The CA model does not have to be totally dismissed, but rather it should be revised to incorporate the idea of typological markedness as the basis for determining degree of difficulty in L2 phonological acquisition (Eckman 1987: 68). This incorporation of typological markedness can be explained via the Markedness Differential Hypothesis (MDH). According to this hypothesis, not all differences between the L1 and L2 will be sources of difficulty. If two structures are different in the two languages but there is no markedness differential between them, the structure will not be difficult to acquire. If the L2 structure is more marked than the L1 structure, it will be more difficult to acquire (Eckman 2004: 530). In this way, Markedness can be seen as a variation on CA in that the relationship between L1 and L2 structures plays a role in the level of difficulty of acquisition.

However, Markedness also depends on linguistic typology, which extends beyond the two languages at the learner’s immediate disposal. The Structural Conformity Hypothesis (SCH) gives attention to the universal aspect of Markedness. It states that IL grammars hold the same universal properties as primary languages (Eckman 2004: 532). IL grammars form a continuum between the L1 and L2, but they always conform to universal generalizations (Eckman 2004: 533). In summary, Markedness is a model that incorporates both L1 and L2 grammars as part of the larger scheme of universal grammar. Within this model learners use more than just the L1 and L2 material in IL phonology, they are able to test universal markedness relationships.
Another demonstration of Markedness’ strength as a model of L2 phonology acquisition is its inclusion in relatively new accounts of SLA done in Optimality Theory, which is the topic of the next chapter (Eckman 2004: 541). Markedness’ strength is in its ability to account for IL forms that are not accounted for by models that only look to the L1 and the L2 as sources of linguistic information. The next section will anticipate difficulties that English speakers will experience while pronouncing Spanish diphthongs based on a comparison between the two languages along with a consideration of markedness relationships.

2.7.2 Pronouncing Spanish Diphthongs

This section begins by anticipating sources of pronunciation errors by comparing and contrasting Spanish and English diphthongs according to the CA model. It then looks to markedness relationships between the diphthongs in each language to further inform the predictions made.

Different Spanish diphthongs present similar pronunciation problems for English speakers. In many cases, Spanish and English diphthongs are comparable, but there are articulatory distinctions that separate them. The Spanish diphthongs [ei], [ai], and [oi] are comparable to English [ei], [ai], and [ɔi], respectively. However, these diphthongs are phonetically distinct between the two languages. In this set of diphthongs, the Spanish glide is faster than in English and the end position for the Spanish diphthong is higher and more forward than in English (Stockwell & Bowen 1965: 98-100). For example, in English eye [ai], the tongue stays on [a] longer. In Spanish hay [ai] ‘there is/are’, the glide is quicker and arrives at a higher position (Whitley 2002: 30). MacLeod (2012) shows that duration is a differentiating factor in L2 diphthongs.
[Vũ] diphthongs present similar problems for English speakers. Spanish [au] is similar to English [ao]. However, this glide in Spanish is tenser, more back, and is pronounced with more rounding than its English counterpart. Spanish [eu] is rare, but presents a problem to English speakers because there is no similar diphthong in English. Both [Vi] and [Vu] diphthongs in Spanish are intelligible when pronounced similarly to their English counterparts, but an L2 accent is perceived by native speakers (Stockwell & Bowen 1965: 98-101).

Spanish has less phonotactic constraints on the distribution of glides than English. GV sequences have a wider distribution after consonants in Spanish than in English. Spanish sequences such as [pu], [ru], [ri], [pi], [mu] are difficult for English speakers, as they rarely, if ever, occur in English. This difficulty prompts pronunciations like [muw-], in which speakers insert the corresponding vowel before the glide, for example [muwerte] for muerte ‘death’ (Whitley 2002: 34). English also has the combination [ju] in words such as music, cute, and beauty (Whitley 2002: 29). While Spanish has the glide [ju] in words such as ciudad ‘city’, its phonetic context in English is limited. Also, the phonemic status of the glide in this sequence is debatable in English, along with its position in the syllable (see section 2.6.2). The glide’s phonemic status should not interfere with the learner’s pronunciation of the Spanish diphthong. However, the limited phonetic contexts for [ju] in English may interfere with pronunciation of the Spanish diphthong with its wider distribution after consonants. MacLeod (2012) confirms that the avoidance of illicit English CjV clusters in L2 Spanish does affect diphthong pronunciation.
The English vowel that most closely resembles Spanish /e/ is commonly pronounced as the diphthong [eɪ], especially in word-final position. This pronunciation is problematic because Spanish has the contrasting diphthong [ei], as in des ‘you give – subjunctive’/ dei ‘you - pl. give – subjunctive’, pena ‘pain’/ peina ‘combs – 3rd person sing.’, reno ‘reindeer’/ reino ‘kingdom’. Spanish /e/ is not pronounced with a glide and it is shorter and tenser (Stockwell & Bowen 1965: 96-97). Hence, the merging of words like those listed above to the diphthongal pronunciation can cause misunderstanding of English speakers (Whitley 2002: 29, 59).

Stockwell & Bowen (1965) propose a hierarchy of difficulty for pedagogical purposes based on the potential for miscommunication caused by improper L2 pronunciation and on the functional load of the segment, which may cause a higher frequency of foreign accent. Their ranking is also based on the congruity between Spanish and English on a given segment (Stockwell & Bowen 1965: 17). At the segmental level, the pronunciation of stressed vowels and diphthongs in Spanish is ranked second on the above-mentioned hierarchy, the pronunciation of unstressed vowels being the most critical aspect of pronunciation on which instructors and students should focus their attention (Stockwell & Bowen 1965: 17, 122).

The reason for a high occurrence of errors and potential for misunderstanding is explained in this way: “Phonological conflicts between the English and Spanish vowel systems depend upon the phonetic realizations of the English phonemes which are incorrectly substituted for Spanish phonemes” (Stockwell & Bowen 1965: 162). Confusion occurs because English has a greater allophonic range for vowels that tolerates multiple pronunciations. The Spanish vowel sounds that the English-speaking student
hears are filtered through the English phonemic inventory and converted to English phonemes. Hence, it is necessary to eliminate the process through which foreign language sounds are filtered through the L1 phonemic system (Stockwell & Bowen 1965: 162-163).

In sum, it should be expected that Spanish diphthongs pronounced by English speakers should have a glide with a longer duration and a lower endpoint than in Spanish. Adjacent consonants may cause pronunciation problems for rising diphthongs, as there are more phonotactic constraints against C+glide clusters in English than in Spanish and phonological processes, such as palatalization (see section 2.3.2), that resolve these clusters. According to a contrastive approach, these pronunciation errors arise because learners are filtering L2 information through their L1 phonology.

These comparisons are made based on the assumption that a diphthong is produced. However, the learner has other options for pronouncing a vowel sequence. Stockwell & Bowen point to consonant, glide, or glottal stop insertion as common strategies for resolving vowel clusters in English (1965: 108-109). The learner could also choose to pronounce the vowels in hiatus, giving the glide of a diphthong full vowel status and increasing the number of syllables. Hiatus creation is more marked than diphthongization, but having only vowels as syllable nuclei is the most unmarked syllable type (Harris 1987: 143, 147). This markedness relationship leaves open the possibility for a hiatus pronunciation as learners may be likely to leave only vowels in nuclear position. Although hiatus is marked, it is also an option in Spanish (see section 2.6.1). Learners may also resolve a vowel sequence by deleting one of the vowels. In order to achieve the maximum rise in sonority as demonstrated by the unmarked CV syllable
(Clements 1990), the less sonorous vowel should be deleted. Comparisons and universal markedness relationships are necessary when predicting L2 pronunciation. With the variety of possibilities available to the L2 learner and the inability of theories of L2 phonology to account for all possible L2 outputs, it should not be surprising that learners may produce forms and use strategies not anticipated here in their attempts to pronounce Spanish diphthongs.

2.8 Conclusion

This chapter has described and compared the Spanish and English vowel systems, syllabification patterns, and the influence of stress on syllabification and vowel pronunciation. In addition to language specific rules, this chapter also considered sonority as a universal property and the role of markedness in shaping syllable structure. The questionable phonemic status of glides in both Spanish and English was also discussed. The chapter then reviewed theories of L2 phonology that attempt to make predictions for the successful acquisition of L2 phones based on comparisons between the L1 and L2 and markedness relationships. The chapter ends by predicting possible L2 pronunciations of Spanish diphthongs by native speakers of English based on the acquisition theories reviewed in this chapter.
Chapter 3: Constraints and Predictions

3.1 Introduction

This chapter begins by describing Spanish and English syllabification within an Optimality Theoretic (OT) framework (Prince & Smolensky 1993/2004), focusing on those aspects that pertain to diphthongization and that affect vowel and glide pronunciation. In an OT account, universal markedness and faithfulness constraints are ranked according to a language specific hierarchy. That is, all languages share a set of universal constraints, but languages differ in how the constraints are ranked. In an OT phonology, surface forms are not derived through ordered rules as in a Sound Pattern of English derivational framework (Chomsky & Halle 1968). Rather, the output form that surfaces is one that has the least violations of the highest ranking constraint. Instead of presenting as exceptions those forms that break phonological rules, OT allows surface forms to violate constraints, and constraints can be re-ranked to account for variation.

This chapter will also consider OT’s role in SLA. While OT has not been used extensively in SLA research, learnability models designed for L1 acquisition, such as the Gradual Learning Algorithm (GLA) (Boersma & Hayes 1999), have applications for SLA that will be reviewed in this chapter. As was discussed in Chapter 2, the principle goal of Contrastive Analysis (as in Stockwell & Bowen 1965) and the Markedness Differential Hypothesis (Eckman 1981) was to make predictions for the relative ease of acquisition of specific elements in the L2 based on L1 and L2 comparisons. This chapter will question how much predictive power OT models have and if OT learnability models are able to overcome the limitations of contrastive models of L2 phonology.
3.2 Optimality Theory

This section outlines the basic tenets of Optimality Theory (Prince & Smolensky 1993/2004). OT is a generative theory of phonology, but it differs from other models in that it is not derivational. In other words, there are no rules to derive surface forms from an underlying representation. Instead of phonological rules, there exist universal constraints that all languages share. Languages differ in how they rank these constraints. Constraints do not function as inviolable rules; rather, all constraints are violable. These violable constraints are arranged in a hierarchy so that the resulting form is the one that incurs the least violations of the highest ranking constraint(s). In other words, the resulting output form is the most optimal candidate; it is the candidate that remains after all the other candidates have been eliminated by violating lower ranked constraints. The winner is not necessarily a perfect candidate, but it is the optimal candidate.

There are two types of constraints in OT. Faithfulness constraints require consistency or identity between input and output forms. The most faithful candidate is the one that deviates least from the input form. Candidates that differ from the input necessarily violate a faithfulness constraint. Two examples of faithfulness constraints are MAX and DEP (McCarthy & Prince 1995). MAX requires that the output have at least as many segments as the input. Thus, deletion of an input segment would violate MAX. DEP requires that there is nothing added in the output that does not exist in the input. For example, epenthesis violates DEP.

Markedness constraints favor unmarked structures. Unmarked structures are those that are more common cross-linguistically. An example of a markedness constraint
is *CODA (NO CODA), which gives a violation mark to output candidates with syllable codas. Syllable codas are generally marked cross-linguistically. Markedness constraints are based only on output forms, whereas faithfulness constraints depend on input forms. Both types of constraints are necessary in any linguistic hierarchy of constraints. If only markedness constraints existed, only very reduced output forms would exist. If only faithfulness constraints existed, there would be no variation between languages. Hence, OT depends on interaction between faithfulness and markedness constraints.

The following sections will discuss the OT literature relevant to syllabification in Spanish and English. Many of the same constraints are active and highly ranked in both languages, but the relative importance of faithfulness and markedness in each language and for each process will determine the language specific rankings that differentiate the two languages.

3.3 Spanish syllable

This section will summarize the relevant constraints and constraint rankings for Spanish syllabification. Colina (2009) offers the most complete account of Spanish syllabification from an OT perspective. The constraint rankings that she proposes for diphthongization and other syllabic processes help to form the basis of the analysis presented in this study. The focus of this section is on diphthongization with attention given to relevant onset and coda constraints that affect vowel and glide pronunciation and where necessary to discuss basic syllable structure. The constraints and the rankings presented in Colina’s work are summarized below.
3.3.1 Basic Constraints

The basic constraints involved in Spanish syllabification appear below in Fig. 3.1. Constraint sets are built upon these basic constraints.

Fig. 3.1 Basic markedness and faithfulness constraints (Colina 2009: 12-13)

<table>
<thead>
<tr>
<th>Basic markedness constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONSET: syllables must have an onset</td>
</tr>
<tr>
<td>*CODA: no coda</td>
</tr>
<tr>
<td>*COMPLEX ONSET: no complex onset</td>
</tr>
<tr>
<td>*COMPLEX CODA: no complex coda</td>
</tr>
<tr>
<td>*COMPLEX NUCLEUS: no complex nucleus</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Basic faithfulness constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX-IO: a segment present in the input must have a corresponding segment in the output (no deletion)</td>
</tr>
<tr>
<td>DEP-IO: a segment present in the output must have a corresponding segment in the input (no epentheses)</td>
</tr>
<tr>
<td>FAITH: umbrella term for all faithfulness constraints</td>
</tr>
<tr>
<td>IDENT: maintain identity of features between the input and output</td>
</tr>
</tbody>
</table>

3.3.2 Onset and Coda

Through the ranking of markedness and faithfulness constraints, OT is able to account for all allowable syllable types in Spanish (see Chapter 2 for allowable syllable constituents). Because Spanish allows syllables with and without onsets and codas, Spanish syllables must violate basic ONSET and CODA constraints. Onsetless syllables violate ONSET, and syllables with a coda violate *CODA. These syllables deviate from the unmarked CV syllables, so that they incur markedness violations. In V and VC syllables, FAITH outranks markedness constraints. High-ranking FAITH also prevents deletion or epentheses and favors more marked syllable types over adding to or taking away information from the input (Colina 2009: 13-14).
Fig. 3.2 Allow onsetless syllables and syllable codas (Colina 2009: 14)

\[
\text{FAITH} \gg \text{ONSET}, *\text{CODA}
\]

\[
\begin{array}{|c|c|}
\hline
\text{\text{o.la}} & * \\
\hline
\text{\text{to.la}} & *! \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|}
\hline
\text{\text{l.a.pis}} & * \\
\hline
\text{\text{l.a.pi.se}} & *! \\
\text{\text{l.a.pi}} & *!
\hline
\end{array}
\]

\text{Sonority}

Sonority constraints determine preferred onset segments, which are those that best conform to the preferred sonority contour by providing a maximum rise in sonority (Clements 1990). Vowels are barred from the onset, as are glides in many dialects, and they are always parsed in the nucleus. Obstruents are the preferred onset, as they are the least sonorous segments. The high ranking of *ONSET/vowel and *ONSET/glide over FAITH blocks a vowel or glide from occupying onset position. Although obstruents are the preferred onset, liquids and nasals are also maintained in onset position due to the dominance of FAITH (Colina 2009: 19-20). The constraint set for sonority in onset position appears below in Fig. 3.3.

\text{Fig. 3.3 Sonority scale: least sonorous segment in the syllable onset} (Colina 2009: 20)

\[
*\text{ONSET/vowel} \gg *\text{ONSET/glide} \gg \text{FAITH} \gg *\text{ONSET/liquid} \gg *\text{ONSET/nasal} \gg *\text{ONSET/obstruent}
\]

Sonority constraints also determine preferred coda segments. The preferred sonority contour calls for a minimal decrease in sonority from the peak (Clements 1990).
High-ranking FAITH favors maintaining an obstruent rather than deleting it, but it is more desirable to have a sonorous consonant in coda position. The constraint set for sonority in coda position appears below in Fig. 3.4.

**Fig. 3.4  Sonority scale: most sonorous segment in coda position**
(Colina 2009: 21)

FAITH >> *CODA/obstruent >> *CODA/nasal >> *CODA/liquid >>
*CODA/glide >> *CODA/vowel

From the constraint ranking in Fig. 3.4 it seems that a vowel is the preferred coda segment. However, because of the universal constraint *CODA and the lack of a *NUCLEUS constraint along with the low ranking of *NUC/vowel (see section 3.3.3 below), the preferred parsing for a vowel will always be in the nucleus over coda position (Colina 2009: 21).

**Complex Onsets and Complex Codas**

Because Spanish allows complex onsets and complex codas, FAITH must dominate *COMPLEX ONSET and *COMPLEX CODA. Again, faithfulness is preferred over deletion (Colina 2009: 14).

**Fig. 3.5  Allow complex onsets and complex codas**

FAITH >> *COMPLEX ONSET, *COMPLEX CODA

As discussed in section 2.3.1, complex codas are rare in Spanish, and coda consonants are susceptible to reduction. The high ranking *COMPLEX CODA, as seen in Fig. 3.6, demonstrates this fact. High-ranking *CODA/stop expresses the undesirability of stops in coda position. The tendency for stop deletion in complex codas is supported by the fact that MAX-IO is ranked below *CODA/stop. Although not commonly occurring (see section 2.3.1), complex codas with a glide + /s/ or glide + /n/ are maintained, as they
follow the sonority contour and are less favored segments for deletion due to their ranking below MAX-IO. As stated in section 3.3.2 and because of other intervening constraints, vowels are the only segments that cannot be parsed in the coda (Colina 2009: 37). Fig. 3.6 shows the revised coda constraint set.

Fig. 3.6 Revised coda contraint set (Colina 2009: 38)

*COMPLEX CODA >> *CODA/stop >> MAX-IO >> *CODA/fricative >> *CODA/nasal >> *CODA/liquid >> *CODA/glide

Onset Processes

This section will discuss onset maximization and onset strengthening.

Onset Maximization

Word-internally, the parsing of consonants in the onset is preferred to the parsing of consonants in the coda in adjacent syllables. This process is known as onset maximization or, in derivational terms, the CV Rule (Hualde 2005: 73). The creation of a complex onset is favored over the existence of a coda consonant; hence, *CODA dominates *COMPLEX ONSET (Colina 2009: 14).

Fig. 3.7 Onset maximization (Colina 2009: 15)

*CODA >> *COMPLEX ONSET

/libro/ libro ‘book’

<table>
<thead>
<tr>
<th></th>
<th>*CODA</th>
<th>*COMPLEX ONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>ℓ♭.ro</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>li♭.ro</td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

Onset Strengthening

Onset strengthening is the process by with a high vocoid becomes a fricative in order to satisfy both ONSET and *ONSET/glide. This process violates a faithfulness constraint of the feature [consonantal]. Strengthening syllable initial glides, thereby
changing their consonantal status, is favorable to an onsetless syllable or the presence of
a glide in onset position (Colina 2009: 24-25).

**Fig. 3.8 IDENT(cons)** (Colina 2009: 24)

IDENT(cons): a segment [α consonantal] in the input must be [α
consonantal] in the output

**Fig. 3.9 Onset strengthening** (Colina 2009: 24-25)

*ONSET/glide, ONSET >> MAX-IΩμ, IDENT(cons)

/kon ielo/ con hielo ‘with ice’

<table>
<thead>
<tr>
<th></th>
<th>*ONSET/glide</th>
<th>ONSET</th>
<th>MAX-IΩμ</th>
<th>IDENT(cons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kon.ye.lo</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>kon.je.lo</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Coda Processes**

As discussed in section 2.3.1, coda consonants often undergo a variety of
neutralization processes or are deleted. Piñeros (2001) presents an OT analysis of
consonant vocalization in Chilean Spanish. Through this process, syllable-final stops
become [u] or [i] and hence, are eliminated from coda position. The motivation behind
vocalization is to achieve harmonious alignment. That is, consonants would rather be
aligned with the left edge of a syllable than with the right, in onset position rather than in
the coda (Piñeros 2001: 164). This analysis will not be detailed here, as the parsing of
post-vocalic glides in the nucleus is different from the parsing in the rest of this analysis.
Postvocalic glides are typically parsed in the coda in Spanish. Nevertheless, Piñeros’
account of Chilean vocalization demonstrates how OT accounts for coda neutralization
strategies in Spanish.5

5 Piñeros also gives an account for consonant deletion in complex codas in Chilean Spanish and for
lateralization of coda consonants in other dialects of Spanish, which demonstrates how OT can handle
3.3.3 Nucleus and Diphthongization

This section will expand on the basic constraint set in Fig. 3.1 concerning constraint ranking in the nucleus and for the process of diphthongization. In order to fully account for glide formation and diphthongization, moraicity and sonority must be considered. The syllable nucleus carries phonological weight, which is described in terms of moras. A vowel has one mora, while glides are nonmoraic. In order to determine which vowel is more susceptible to gliding, it is necessary to consider sonority. Low vowels are the most sonorous and are more likely to maintain their mora. Conversely, high vowels are least sonorous and more likely to be nonmoraic. In other words, high glides are more desirable than low glides. The constraint that assigns a mora to a high vowel is the most likely to be violated in the constraint set in Fig. 3.10. In fact, /a/, which is never never pronounced as a glide, must be moraic, while all other Spanish vowels can be nonmoraic (Colina 2009: 26).

Fig. 3.10 Glide formation in relation to vowel height (Colina 2009: 26)

\[
\begin{align*}
\text{low} /\mu & >> \text{mid vowel} /\mu >> \text{ONSET} >> \text{hi} /\mu \quad (\text{only high glides}) \\
\text{low} /\mu & >> \text{ONSET} >> \text{mid vowel} /\mu >> \text{hi} /\mu \quad (\text{only high and mid glides})
\end{align*}
\]

In rising diphthongs, prevocalic glides are parsed as part of the nucleus. This parsing creates a complex nucleus, thus violating *COMPLEX NUCLEUS and *NUC/glide (Colina 2009: 21-22). While a vowel can be the only single segment in the nucleus, a glide may be present when accompanied by a vowel. The constraint set below demonstrates that a glide is a permissible segment in the nucleus due to the ranking of FAITH over *NUC/glide. A vowel is the best possible nucleus, but a glide may accompany the vowel as the next sonorant element on the sonority scale. The high

**Fig. 3.11 Sonority scale: most sonorant segment in the syllable nucleus**
(Colina 2009: 20)

*NUC/obstruent >> *NUC/nasal >> *NUC/liquid >> FAITH >> *NUC/glide >> *NUC/vowel.

As discussed in section 3.3.2 and as can be seen in Fig. 3.3, vowels are barred from the onset, as are glides in many dialects, and are always parsed in the nucleus. The high ranking of *ONSET/vowel and *ONSET/glide over FAITH blocks a vowel or glide from occupying onset position and maintains their preferred position in the nucleus. A complete ranking must also concern syllable weight. A vowel gives up its mora during diphthongization in order to maintain an allowable syllable weight. A low ranking faithfulness constraint that obligates mora correspondence in the input and output is violated during diphthongization (Colina 2009: 22-23).

**Fig. 3.12 MAX-IOμ (Colina 2009: 23)**

MAX-IOμ: a mora present in the input must have a corresponding mora in the output

**Fig. 3.13 Diphthongization (rising diphthongs only) (Colina 2009: 23)**

*ONSET/glide, ONSET >> *COMPLEX NUCLEUS, *NUC/glide, MAX-IOμ

/niebe/ nieve ‘snow’

<table>
<thead>
<tr>
<th></th>
<th>*ONSET/glide</th>
<th>ONSET</th>
<th>*COMPLEX NUCLEUS</th>
<th>*NUC/glide</th>
<th>MAX-IOμ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>*nje.βe</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b.</td>
<td>ni.e.βe</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Postvocalic Glides

Postvocalic glides are parsed in the coda, thereby violating *CODA or *COMPLEX CODA (Colina 2009: 18). Diphthongization, and thereby coda creation, is preferable to hiatus, which would result in a violation of high ranking ONSET.

**Fig. 3.14 Diphthongization (falling diphthongs)** (Colina 2009: 42)

\[ *\text{ONSET/glide}, \text{ONSET} \gg *\text{COMPLEX NUCLEUS}, *\text{CODA/glide}, *\text{CODA}, \text{MAX-IO} \]

/pausal pausa ‘pause’

<table>
<thead>
<tr>
<th></th>
<th>*ONSET/glide</th>
<th>ONSET</th>
<th>*COMPLEX NUCLEUS</th>
<th>*CODA/glide</th>
<th>*CODA</th>
<th>MAX-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>“pau.sa”</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>“pa.u.sa”</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3.4 Syllabic Processes

The following section discusses resyllabification and diphthongization across words.

**Resyllabification**

As in section 2.5.1, resyllabification is the process in which a word-final consonant is parsed, or resyllabified, as the onset of a following word in order to avoid an onsetless syllable. In OT, resyllabification is accomplished via alignment constraints.

**Fig. 3.15 Alignment constraint for resyllabification** (Colina 2009: 48)

ALIGN-LST: align the left edge of the stem with the left edge of the syllable

In resyllabification, ALIGN is violated in favor of satisfying ONSET. However, resyllabification does not create complex onsets, which means that ALIGN-LST must dominate *CODA. In a derivational approach, it is said that resyllabification is a post-
lexical process, necessitating ordered rules. In an OT approach, ALIGN-LST’s dominance over *CODA demonstrates the preference for proper alignment of coda consonants when ONSET is already satisfied. The motivation for resyllabification, which is onset maximization, is captured in an OT approach (Colina 2009: 48-49). The constraint hierarchy for resyllabification appears below.

**Fig. 3.16 Constraint ranking for resyllabification** (Colina 2009: 49)

\[
\text{ONSET} >> \text{ALIGN-LST} >> \text{*CODA}
\]

<table>
<thead>
<tr>
<th>/mis amigos/ mis amigos ‘my friends’</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \overset{\circ}{\text{ONSET}} )</td>
</tr>
<tr>
<td>( \overset{\circ}{\text{mi.s}</td>
</tr>
<tr>
<td>( \overset{\circ}{\text{mis.a.mi.,\gamma}} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/kon ielo/ con hielo ‘with ice’</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \overset{\circ}{\text{ONSET}} )</td>
</tr>
<tr>
<td>( \overset{\circ}{\text{kon.ye.lo}} )</td>
</tr>
<tr>
<td>( \overset{\circ}{\text{ko.nje.lo}} )</td>
</tr>
</tbody>
</table>

**Resyllabification and Diphthongization**

Just as diphthongization occurs within words, it also occurs across words with the same motivation to avoid ONSET violations. Hence, ONSET dominates both ALIGN-LST and MAX-IO\( \mu \), as shown below (Colina 2009: 52).

**Fig. 3.17 Diphthongization across words** (Colina 2009: 52)

\[
\text{ONSET} >> \text{ALIGN-LST}, \text{MAX-IO}\mu
\]

<table>
<thead>
<tr>
<th>/tu amigo/ tu amigo ‘your friend’</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \overset{\circ}{\text{ONSET}} )</td>
</tr>
<tr>
<td>( \overset{\circ}{\text{tu</td>
</tr>
<tr>
<td>( \overset{\circ}{\text{tu.a.mi.,yo}} )</td>
</tr>
</tbody>
</table>

It is possible that L2 learners will adopt the same strategies for resolving hiatus that native speakers use. It is also possible that English constraint hierarchies will play a
role in accounting for the pronunciation of diphthongs in L2 Spanish. The following sections will discuss constraint hierarchies for English syllabification, especially those having to do with vowel and glide pronunciation in English.

3.4 The English Syllable

This section will discuss the syllable in English in an OT framework. The information in the following sections is based on Hammond (1999) and Hall (2001, 2006). While Hammond (1999) offers an OT account of English phonology that is comparable to Colina (2009) in its completeness, many of his arguments are counter to widely held claims of universal markedness. Hall (2006) offers an alternative analysis of English syllabification that is in line with markedness universals and that finds little need for language specific constraints.

This section will highlight the roles of moraicity and sonority in determining permissible constituents for each part of the syllable in an OT framework. While stress plays an important role in English syllabification and vowel pronunciation, an analysis of stress is not the focus of this study and will only be included where necessary.

3.4.1 Moraicity

In both Hammond (1999) and Hall (2001), moraicity is central to the discussion on English syllabification. English content words must be minimally bimoraic (Hammond 1999: 41) and maximally trimoraic. The minimum and maximum limits are stated in the following constraints, which are based on Hammond (1999).

Fig. 3.18 Bimoraic minimum word constraint (Hammond 1999: 135)

BIMORAICITY – words must contain at least two moras
Fig. 3.19 Trimoraic maximum syllable constraint (Hammond 1999: 136; Hall 2001:406)

3/μ – syllables must contain no more than three moras

Both BIMORATICITY and 3/μ must outrank FAITH, as segments may be deleted or modified to fit the two to three mora window (Hammond 1999: 136). These requirements place limits on what segments can pertain to different syllabic positions.

Onset

Mora count is irrelevant in the onset, as the onset does not affect syllable weight (Hammond 1999: 41). There can be as many segments as the language will allow in onset position, which is three in English (see section 2.3.2). The preference for an onset is stated in the following constraint.

Fig. 3.20 Basic onset constraint (Hammond 1999: 133)

ONSET – syllables must have onsets

Coda

The contribution of coda consonants to mora count varies. This variability has to do with the moraic status of coronal consonants. In certain circumstances, which will be discussed below, coronal consonants may contribute one mora in coda position or none at all. However, noncoronals must contribute a mora in coda position. According to Hammond, there are also coda consonants that add two moras (1999: 105). As in Hammond, Hall’s analysis allows for coronals to be unassociated with a mora; however, all other consonants contribute no more than one mora (Hall 2001). The following section presents two distinct analyses of the syllabification of word medial consonants.
These analyses demonstrate differing interactions between onset and coda constraints, along with differing accounts for the optional moraicity of coronal codas.

**Word-Medial Consonants**

The parsing of word-medial consonant clusters is a source of controversy in English syllabification. Hammond (1999) and Hall (2006) propose distinct OT analyses for the parsing of these clusters. Hammond accounts for the distribution of vowels and consonants in English by proposing that word medial single consonants must affiliate to the right (V.CV) and that word medial coda clusters must affiliate to the left (VCC.V). He calls for coda maximization in the parsing of word medial consonant clusters (Hammond 1999: 133-134). Hall (2006) opts for onset maximization. His analysis would agree with Hammond’s parsing of single consonants to the right (V.CV); however, his analysis of word medial consonant clusters would affiliate at least one consonant to the right (VC.CV or V.CCV) (Hall 2006: 7). Hall’s analysis is better motivated, considering the universal preference for CV syllables.

To begin, Hall (2006) proposes an onset well-formedness constraint. This constraint calls for onsets to be well-formed according to the Sonority Sequencing Generalization (SSG) and according to individual constraints that limit permissible onset sequences.

**Fig. 3.21 Onset well-formedness constraint**

\[ \text{OWF} \Rightarrow \text{onset clusters must adhere to SSG and to the individual constraints } *_o[t, *_o[t[n, *_o[n] \]

Any resulting coda clusters also must conform to the SSG.
Fig. 3.22 Coda well-formedness constraint

SSG-CODA -- coda clusters must follow SSG

The onset and coda well-formedness constraints are undominated with respect to each other (Hall 2006: 6).

Hall’s analysis also makes use of the basic universal constraints NOCODA, NOCOMPCODA, and NOCOMPONSET. The ranking below accounts for the correct parsing of VCCV sequences into V.CCV or VC.CV sequences as necessary to satisfy syllable well-formedness.

Fig. 3.23 Constraint ranking for VCCV > V.CCV

NOCODA >> NOCOMPONSET

<table>
<thead>
<tr>
<th>/distant/ distant</th>
<th>NOCODA</th>
<th>NOCOMPONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>dist.nt</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>dis.tnt</td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>dr.stnt</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

This ranking shows the preference for onset creation over coda creation, even if the resulting structure is a more marked complex onset. The optimal form ends in a lax vowel, but it is still preferred over forms that heterosyllabify the sC cluster. This ranking also accounts for the fact that the stop is not aspirated as it would be in initial onset position (Hall 2006: 10-11). While the optimal candidate does not conform to the OWF constraint, sC clusters are a well-known exception to the onset sonority profile and are acceptable onsets in English.

VCCV sequences are parsed as VC.CV sequences to avoid violations of OWF and to avoid a complex coda. The constraints NOCOMPCODA and OWF are added to the above constraint ranking.
Fig. 3.24 Constraint ranking for VCCV > VC.CV (Hall 2006: 12)

NOCOMP CODA, OWF >> NOCODA >> NOCOMPONSET

<table>
<thead>
<tr>
<th>/welkam/ welcome</th>
<th>NOCOMP CODA</th>
<th>OWF</th>
<th>NOCODA</th>
<th>NOCOMPONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>welk.am</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>œwel.kam</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>wel.kam</td>
<td>*!</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Because [lk] is not a permissible onset cluster and because a complex onset is undesirable, the consonant cluster is split to create an onset at the expense of maintaining a coda consonant. The optimal candidates in Fig. 3.34 and 3.35 have word-medial onsets, demonstrating the motivation for onset maximization in the data up to this point.6

The constraint ranking above is amended to account for the correct parsing of VC(CC)GV sequences. Hall parses all prevocalic glides as part of the onset, which is different from other analyses, such as Baertsch (2008), which will be discussed in section 3.4.2. Below is the constraint ranking to account for VC(CC)GV sequences.

Fig. 3.25 Constraint ranking for VC(CC)GV sequences
(Hall 2006: 19-22)

ONSET, OWF >> NOCOMP CODA >> NOCODA >> NOCOMPONSET

<table>
<thead>
<tr>
<th>/milion/ million</th>
<th>ONSET</th>
<th>OWF</th>
<th>NOCOMP CODA</th>
<th>NOCODA</th>
<th>NOCOMPONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>œmil.jn</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mil.jn</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/œkjumen/ acumen</th>
<th>ONSET</th>
<th>OWF</th>
<th>NOCOMP CODA</th>
<th>NOCODA</th>
<th>NOCOMPONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>œk.jumŋ</td>
<td></td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>œœ.kjumŋ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

6 Hall also presents an analysis of VCCCV and VCCCCV sequences which will not be reviewed here. The analysis of these sequences further serves to highlight onset maximization as the motivation for parsing word-medial consonant clusters. See Hall (2006: 15-19) for the analysis of these sequences.
CG clusters that are not permissible onsets in English, such as [lj] and [nj], would be eliminated as onset clusters by OWF and parsed heterosyllabically. However, all permissible stop + G clusters would be parsed as onset clusters.

Hall’s analysis has the benefit of concurring with universal markedness principles. He would parse a word like *matron* as [meɪ.ən], thereby maximizing the onset and eliminating a coda when possible. Hammond’s analysis favors coda maximization, which runs counter to principles of universal markedness and would parse *matron* as [mæt.ən]. However, Hammond’s analysis considers moraicity and takes vowel quality into account, whereas Hall’s analysis make little mention of vowel quality and does not consider the effect of moraicity on syllabification.

**Nucleus**

Both Hammond and Hall’s accounts of English syllabification rely on ONSET and CODA constraints. NUCLEUS constraints do not play the same role that they do in Colina’s account of Spanish syllabification. The discussion on the parsing of word-medial consonants is lengthy in both English accounts, but the discussion as to what is allowed in the nucleus is lacking.

As stated in section 2.3.2, English has lax and tense vowels, and these vowels have different weights. Hammond’s mora count in the nucleus is as follows. “Lax vowels have one mora, tense vowels and the diphthongs [ay, yu] have two moras, and the diphthongs [aw, øy] have three moras” (Hammond 1999: 135). Because lax vowels contribute one mora, monosyllabic words cannot end in lax vowels as such a word would violate BIMORAICITY. Tense vowels have two moras, so that monosyllabic words ending in tense vowels are permissible (Hammond 1999: 135). Hall does not make a distinction
between the moraicity of different diphthongs. In his account, all diphthongs are bimoraic (Hall 2001: 405). The discrepancy in mora count in the two analyses will be discussed below in section 3.4.2.

Concerning the falling diphthongs, Hammond states that the “Consonants [w,y] can only occur word-finally as part of the diphthongs [aw, ay, oy]” (1999: 34). He further states that “The restrictions against [w, y] in coda position except when part of a diphthong would be easily treated if the glide portion of the diphthongs were not really in coda position. If this can be maintained, constraints like *CODA/w and *CODA/y would do the trick” (Hammond 1999: 45). According to this statement, it is unclear whether the glide in a falling diphthong is parsed in the coda or the nucleus. This statement is also unclear as to whether the constraints *CODA/w and *CODA/y are in fact being proposed as active in English.

Concerning rising diphthongs, Hammond states that any vowel can follow any onset consonant with the exception of the diphthong [yu], which cannot follow coronal consonants (1999: 107-108). However, [yu] can follow a coronal sonorant word-medially, as in value [vælju] (Hammond 1999: 120). Hammond parses [yu] in the peak based on the fact that speakers rarely split the segments when asked to do so in word games, demonstrating that the segments pertain to the same part of the syllable (1999: 244-245). Hall parses prevocalic glides as part of the onset but does not provide an explanation to why the glide is parsed in the onset and not in the nucleus (Hall 2006: 19-27). The parsing of prevocalic glides will be further discussed in section 3.4.2.
Summary

There are differences in the ways that Hammond (1999) and Hall (2001, 2006) treat syllable weight in their analyses. The moraicity of diphthongs is different in each account. Hammond (1999) counts the diphthongs [aʊ, ɔɪ] as trimoraic, while Hall (2001) counts all long vowels and diphthongs as bimoraic. This difference in mora count for diphthongs relates to the treatment of coronal coda consonants, which will not be discussed here. Hall (2006) has little mention of mora and instead focuses on the role of universal markedness in syllabification.

In comparing syllable weight in Spanish and English, it can be said that there is less variation in mora count among vowels in Spanish, as there is no distinction between long/short or tense/lax vowels. In Spanish, moraicity is used to account for diphthongization, as vowels lose a mora by becoming glides (Colina 2009). In English, diphthongs have at least the same mora count as long vowels (Hammond 1999; Hall 2001, 2006). The mora count of certain segments in an OT analysis of English syllabification is still debatable, but it is clear that moraicity is a central and necessary element in the discussion.

3.4.2 Sonority

This section will review how OT formalizes sonority’s role in English syllabification. Sonority was discussed in Chapter 2 based on Clements (1990). Hammond (1999) presents more detailed constraint rankings for sonority than in Hall (2006). While the exposition of Hall’s sonority based constraints is more concise and straightforward, Hammond’s is more explicit. This section will discuss the role of
sonority in each part of the syllable and also how OT accounts for coronal consonants that provide exceptions to the preferred sonority contour.

Onset

Hammond’s constraint ranking for sonority in onset position appears below.

**Fig. 3.26 Constraint ranking for sonority hierarchy for onset(o) consonants**  
(Hammond 1999: 87)

\[
\text{FAITH} \gg \text{*ONSET/approximant(a)} \gg \text{*ONSET/nasal(n)} \gg \text{*ONSET/obstruent(o)}
\]

This ranking demonstrates that an obstruent is the preferred onset consonant, providing the greatest rise in sonority towards the peak. An approximant is the least desirable onset, since it creates the least dramatic rise in sonority. Because of high-ranking FAITH, any onset is favored over no onset (Hammond 1999: 87).

Hall (2001) cites Selkirk’s Sonority Sequencing Generalization (SSG) which states that, “In any syllable, there is a segment constituting the syllable peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values” (Selkirk 1984:116). Hall incorporates the SSG into the ONSET WELL-FORMEDNESS (OWF) constraint (see section 3.4.1), stating that sonority along with other constraints will determine permissible onsets. Sonority alone is not sufficient for determining allowable onsets, as some clusters that would be well-formed based on sonority are not acceptable onset clusters in English. For example, \(\sigma [pl]\) and \(\sigma [kl]\) are permissible, but \(\sigma [tl]\) is not (Hall 2006: 6).

Coda

Hammond’s constraint ranking for coda consonants mirrors the onset sonority constraint ranking.
Fig. 3.27 Constraint ranking for sonority hierarchy for coda(c) consonants
(Hammond 1999: 91)

*Coda/obstruent(o) >> *Coda/nasal(n) >> *Coda/l >> *Coda/r

This ranking demonstrates that an obstruent is the least desirable coda, and [r] is the most desirable coda consonant. The most desirable coda provides a minimal drop in sonority (Hammond 1999: 91).

Hall offers the single constraint SSG-CODA, which states that coda clusters conform to the SSG, to account for the role of sonority in determining permissible coda clusters (2006: 6).

While there are exceptions to sonority hierarchy, they will not be discussed here, as these constraint rankings have not impact on the analysis offered in this study. These studies have to do with [s] and other coronal consonants, and are discussed in Hammond (1999) and Hall (2001, 2006).

Nucleus

As has already been stated, the nucleus is the most sonorous element of the syllable. There is some disagreement as to how prevocalic glides are parsed in the onset or the nucleus (see section 3.4.1). Baertsch (2008) offers an OT analysis of underlying CiV and CuV clusters in which she explains their asymmetrical inventories and parsings. In CiV clusters, the glide is parsed as part of the syllable nucleus, forming a diphthong. For example, music /miusɪk/ is pronounced [mju.zək]. The following vowel is always /u/ (Baertsch 2008: 225, 233). In conjunction with any other vowel, the /i/ becomes the nucleus, and a hiatus is formed. For example, tiara /tiːərə/ is pronounced [tʰi.əə] (Baertsch 2008: 235).
Baertsch’s account incorporates the principle of minimum sonority distance and onset and nucleus well-formedness constraints. She uses the Peak and Margin Hierarchies (Prince & Smolensky 1993/2004: 159-160) to account for the role of sonority in the onset and nucleus. She adds a second Margin hierarchy to account for the second segment in the onset. The Peak and Margin hierarchies are as follows.

**Fig. 3.28 Peak and margin hierarchies** (Baertsch 2008: 228 -- based on Prince & Smolensky 1993/2004)

**Peak hierarchy**
*P/Obs >> *P/Nas >> *P/l >> *P/r >> *P/u >> *P/i >> *P/[hi]

**Margin hierarchy**
*M$_1$/hi >> *M$_1$/i >> *M$_1$/u >> *M$_1$/r >> *M$_1$/l >> *M$_1$/Nas >> *M$_1$/Obs

Baertsch’s M$_2$ hierarchy
*M$_2$/Obs >> *M$_2$/Nas >> *M$_2$/l >> *M$_2$/r >> *M$_2$/u >> *M$_2$/i >> *M$_2$/[hi]

The optimal onset, that with the steepest rise in sonority, will violate the lowest ranked constraints. These constraints interact to function as minimum sonority distance parameters. Their interaction with Faithfulness constraints will determine permissible onset clusters in a given language (Baertsch 2008: 228-229).

The constraint ranking for the realization of CuV clusters appears in the following tableau.
Fig. 3.29 CuV sequences (Baertsch 2008: 231)\(^7\)

<table>
<thead>
<tr>
<th></th>
<th>DEP</th>
<th>*M(_1/u)</th>
<th>*P/u</th>
<th>*Ob(_5)u(_2)</th>
<th>*P/i</th>
<th>ONSET</th>
<th>*M(_2/i)</th>
<th>UNIFORMITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{\textsuperscript{c}})k(_1)w(_2)k</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k(_1)wik</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k(_1)ijk</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k(_1)u.ik</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k(_1)ə.w(_1)k</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The optimal candidate parses the underlying /u/ as the second member of the onset after an obstruent, providing a steep rise in sonority. Parsing the /u/ in the peak is less favorable and results in the elimination of three of the candidates. The last candidate in the tableau separates the word into two syllables and parses the /u/ as the first segment of the onset of the second syllable. This parsing is achieved through epenthesis and is eliminated by high ranking DEP. This constraint ranking demonstrates the preference in English for underlying /u/ to be parsed as the second member of an onset cluster with obstruents. Because *P/u is dominated by DEP and NUC, in CuC sequences the /u/ will be parsed in the peak, as it is necessary for a syllable to have a nucleus (Baertsch 2008: 232-233).

The constraint ranking for the realization of CiV clusters appears in the following tableau.

\(^7\) This and the following tableaux based on Baertsch (2008) do not maintain the same numbering of margin segments for the output candidates. Instead, only the margin segments that are of interest to the present research are numbered in subscript.
The optimal candidate parses the /i/ as part of the nucleus. Parsing the /i/ as the second member of an onset cluster is eliminated by ranking the constraint that prohibits obstruent + j clusters above the constraint that disallows /i/ in the peak. In the penultimate candidate, ONSET surfaces to rule out hiatus. Finally, due to high ranking DEP, epenthesis is ruled out, which would have pushed the /i/ to be a glide in onset position, as proposed by the last candidate (Baertsch 2008: 234).

Sequences in which underlying /i/ is optimally parsed as an onglide in the nucleus are always Ciu sequences. When Ci is followed by a vowel other than /u/, the /i/ is parsed as the sole segment in the nucleus, and hiatus is formed. All of the words that fit this pattern are borrowed. The constraint ranking for such sequences is as follows.

**Fig. 3.30 CiV sequences** (Baertsch 2008: 234)

<table>
<thead>
<tr>
<th></th>
<th>DEP</th>
<th>*P/u</th>
<th>*Obs_i2</th>
<th>*P/i</th>
<th>*M_1/i</th>
<th>ONSET</th>
<th>*M_2/u</th>
</tr>
</thead>
<tbody>
<tr>
<td>/piuk/ puke</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p_i_juk</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p_i_2juk</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p_i_2j_u</td>
<td>!</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The optimal candidate parses the /i/ as part of the nucleus and creates a hiatus. An AGREE constraint, which requires that nuclear segments share the feature [+hi], eliminates the candidate with a diphthong in the nucleus because the segments are not both [+hi]. The
constraint against obstruent + j clusters eliminates the candidate that parses the /i/ as a glide and as the second member of an onset cluster (Baertsch 2008: 235-236).

/ɪ/ and /u/ pattern differently as the middle segment in CVV sequences. /ɪ/ “is pulled into nucleus of the syllable”, either as part of the diphthong [ju] or as a simple nucleus (Baetsch 2008: 237). /u/ is parsed as part of an onset cluster when it is followed by a vowel and in the nucleus when it appears before a consonant (Baertsch 2008: 237). Baertsch demonstrates the role of sonority in the parsing of prevocalic glides, and her analysis follows from established OT margin and peak hierarchies. Hall’s (2001, 2006) parses pre-vocalic /i/ in the onset, and his analysis also follows universal sonority and markedness principles. For the purposes of the analysis of L2 diphthong in this study, the parsing of the glide into a particular syllabic constituent is not of utmost importance. What is important is that many CɪV sequences are illicit in English, and Baertsch’s constraint hierarchies are able to capture this sequential relationship between the segments. Baertsch’s (2008) constraint rankings will be used in this study to account for L2 diphthongs.

3.4.3 Stress and Vowel Quality

As stated in Chapter 2, stress and vowel pronunciation are linked in English. [ə] is the most common vowel in English, but it is not a phoneme. Instead, it surfaces when vowels are in an unstressed syllable. Hence, schwa is always derived, and Hammond proposes that it does not bear a mora. In general, unstressed preconsonantal vowels surface as schwa. In word-final position, the high vowels and mid back vowels surface as tense and are bimoraic, while all other vowels surface as schwa, for example in the words pretty [pʰɪɾi] and igloo [ɪɡlu] (Hammond 1999: 205-206). For this situation to
occur, Hammond offers an analysis of ambisyllabic ity, which is not discussed in this study. Also, because of his use of MAX CODA, which is cross-linguistically unmotivated, his analysis of the relationship between stress and vowel quality will not be discussed here.

Instead, Burzio (2007) offers an analysis of this relationship based on perceptual distance. His analysis is able to explain why all unstressed vowels do not reduce. It is also able to account for cross-linguistic differences in the relationship between stress and vowel pronunciation. Although his analysis breaks with the syllabic perspective adopted in the rest of this study, it is an attractive analysis for the reasons mentioned above. Burzio’s (2007) analysis will be discussed in more detail in Chapter 6.

3.4.4 Summary

The preceding section has reviewed aspects of English syllabification relevant to this research. Both moraicity and sonority play a central role in OT accounts of English syllabification. In the case of intervocalic consonant clusters, onset maximization is the preferred motivation as it is more common cross-linguistically. The parsing of prevocalic glides has also been debated. Based on the sections on moraicity and sonority above, it is evident that the Optimality-Theoretic literature on English syllabification has focused less on vocalic processes than has the OT literature on Spanish. The parsing of segments in Spanish is more straightforward and there is more consensus on the topic than in English.

The lack of consensus concerning aspects of English syllabification may provide difficulties for the L2 researcher, as it is unclear which constraint ranking to select as a source of linguistic transfer. The following section will discuss OT research in the field
of SLA. One of the central issues in this discussion will be to determine the extent to which L1 constraint rankings impact L2 acquisition.

3.5 OT and SLA

This section will begin by reviewing the Gradual Learning Algorithm (Boersma & Hayes 1999), highlighting its strengths and discussing its implications for SLA research. The section will also review other Optimality-Theoretic work in SLA.

3.5.1 Gradual Learning Algorithm

The Gradual Learning Algorithm (GLA) of Boersma & Hayes (1999) is an OT learnability model that was designed to account for L1 acquisition. The GLA shares the basic concepts of Tesar and Smolensky’s (1996) error driven Constraint Demotion model, but differs in the strictness of constraint domination. Boersma and Hayes’ (1999) model is still error driven, in that the algorithm only changes constraint rankings when the input data does not agree with the learner’s current ranking hypothesis. However, the GLA differs from the Constraint Demotion model in that it has a continuous scale of constraint strictness instead of adhering to strict domination of constraints. In this continuum of strictness and laxness, higher ranked constraints are stricter and lower ranked constraints are more lax. Furthermore, the grammar is stochastic. At every evaluation point of a candidate, there is a noise component that allows for the grammar to produce variation in the outputs, especially if the rankings are close. That is, the constraints have ranges. The constraint ranges can overlap, and this overlap allows for free variation in output forms. The distance between constraints represents a probability that makes predictions for the frequency of output forms that show variation.
In the GLA, as in Constraint Demotion, the learner starts out at an initial state with a particular set of constraint rankings. After receiving adult input that is considered to be correct and after accessing underlying forms that the learner already has access to as part of a universal inventory, the learner compares the forms. If the forms match, no changes are necessary. If the forms differ, then the learner adjusts the constraint ranking. The learner eventually reaches a final state in which the algorithm repeats until the optimal candidate is generated. The GLA can accommodate optionality and free variation. Speech errors do not present a problem nor do they mislead the learner because of their low frequency. The GLA also allows for intermediate levels of well-formedness because of the continuous scale of constraint strictness (Boersma and Hayes 1999).

3.5.2 Testing the GLA in L1 Acquisition

The GLA has been tested in L1 acquisition research and has been shown to be an adequate and useful model in explaining acquisition processes. Boersma & Levelt (1999) show that the GLA is able to realistically predict acquisition order of L1 syllable structure. They tested the algorithm with a computer simulated learner and actual Dutch speaking children. The authors found that the GLA realistically modeled both fixed and variable order of the acquisition of different syllable types. They also confirm that learning is gradual; it occurs rapidly at first and then slowly rises to complete correctness (Boersma & Levelt 1999: 8). This study shows that the GLA is able to account for developments that are observed in actual L1 acquisition.

Levelt & van de Vijver (2004) also discuss L1 syllable acquisition, but from a cross-linguistic approach. The authors remark on general agreement in the field that at
first markedness constraints rank higher than faithfulness constraints and that by demoting markedness constraints, outputs become increasingly marked and increasingly faithful to their inputs. They also acknowledge the assumption that initial output is generally unmarked and that the marked elements of language are those that the learner needs to acquire. The authors confirm generalizations that are central to the GLA.

The authors also find that a learner’s intermediate grammar will represent the final state grammar of another world language, as demonstrated through the acquisition of syllable structure. For example, if a language allows violation of the constraints NO-CODA and ONSET, at some intermediate point of acquisition the learner’s grammar will represent that of another language that does not allow violation of NO-CODA or of another that does not allow violation of ONSET (Levelt & van de Vijver 2004: 205). By analyzing the development of syllable acquisition cross-linguistically, the authors found that there is not that much developmental difference across languages and that learners of different languages follow similar learning paths, which are determined by the frequency of syllable types. These cross-linguistic observations support the GLA in that frequency of the observed structure determines the amount of variation. If there are close frequencies of different syllable types, there will be more variation in the learning path. The authors combined acquisition and typological data to demonstrate cross-linguistic development of syllable types that confirms many currently held theories about learnability in OT. This study provides cross-linguistic evidence that can be used in support of the GLA in its observations having to do with learner variation.
3.5.3 Testing the GLA in SLA

The GLA has also been tested in SLA research. A preliminary question that must be considered in evaluating learnability models is how L1 and L2 acquisition differ. One difference is that L2 learnability models must consider the role of L1 transfer. In the context of adult SLA, they must also recognize a different initial state from that of L1 acquisition.

Hancin-Bhatt & Bhatt (1997) investigate the role of L1 transfer and developmental effects in the acquisition of L2 syllable structure. The authors consider the differences between L1 and L2 acquisition, but do not test any particular learnability model. The focus of the study is the acquisition of English complex onsets and complex codas by native speakers of Spanish and Japanese. The authors argue that OT can best explain the interaction between developmental and transfer effects in SLA. Developmental effects are universal in nature and make some structures easier to learn than others due to linguistic complexity (Hancin-Bhatt & Bhatt 1997: 332). During the acquisition process, learners do not have to learn new constraints for the L2 because these constraints are universal. Rather, learners must re-rank constraints, mapping L1 constraints onto the L2. The authors posit that EVAL is L1-based for a beginning L2 learner. Over time, the learner rearranges the constraint ranking, but there may still be remnants of the L1 as the constraints resist reorganization (Hancin-Bhatt & Bhatt 1997: 359). Evidence for both non language specific developmental effects and transfer effects were found. Developmental effects may surface as a response to a lack of positive transfer. The learners seem to start off with an initial L1 constraint ranking, but then
move to one that more closely resembles the L2 (Hancin-Bhatt & Bhatt 1997:369). It is evident that the interlanguage (IL) phonology is “dynamic” (Hancin-Bhatt & Bhatt 1997:370) and that OT is able to explain variation in learner performance. The participants used different strategies for pronouncing coda clusters. This variation is due to language-specific constraints.

Swanson (2001) demonstrates how L2 acquisition would proceed through constraint ranking in specific stages of L2 acquisition of English and Polish. While she tests Constraint Demotion rather than the GLA, her findings still apply to the discussion on the GLA. Swanson finds that L1 transfer plays a role in SLA, and that learning L2 phonology can sometimes go through the same process as L1 learning (2001: 23). Learners transfer from the L1, but they also have access to universal grammar. How the initial state is defined is what differentiates the L1 and L2 learning processes. The initial state for L2 learners is their final L1 state (Swanson 2001: 27). In L2 acquisition, learners have to demote faithfulness constraints, whereas in L1 learning, markedness constraints are demoted (Swanson 2001: 33). Tesar and Smolensky’s (1996) model can be used to account for L2 learning, but a modification is necessary. That modification is that learners must demote both faithfulness and markedness constraints. If faithfulness constraints are not demoted in L2 acquisition, there is the possibility for fossilization to occur (Swanson 2001: 35). Defining a final state grammar for the L1 is problematic, as language is dynamic. However, it is reasonable that adult L2 learners will begin L2 acquisition with their current L1 grammar.

Hancin-Bhatt & Bhatt (1997) and Swanson (2001) refer to the learner’s access to universal grammar in addition to the constraints already acquired in the L1. Broselow,
Chen & Wang (1998) demonstrate how universal grammar can be accessed in their focus on the emergence of the unmarked in SLA. The “emergence of the unmarked” refers to a situation in which low-ranked markedness constraints surface to become active under certain circumstances (McCarthy & Prince 1994). Previous research in SLA has found patterns in L2 speech that do not appear in the L1 or L2. These patterns are a result of L2 learners favoring universally less marked structures. Hence, it is possible that low ranked markedness constraints that do not surface in the L1 or the L2 can become active in the IL. Broselow, Chen & Wang posit that using a set of ranked universal constraints, not just those that appear in the L1 and L2, to describe IL grammars can explain learners’ simplification strategies and the choices that learners make about which strategy to use (1998: 263). Because GEN provides universal candidates that go through a hierarchy of universal constraints, there is nothing in the IL grammar that is not part of the L1 or L2. All constraints exist in all languages, but they are not always active in all languages. The authors do agree with other research in that learners start with L1 constraints and work their way towards a ranking more like the L2 (Broselow, Chen & Wang 1998: 269). Because the ranking of IL constraints is “in flux”, markedness constraints that are not visible in the L1 or L2 may have the opportunity to become visible (Broselow, Chen, & Wang 1998: 274).

Broselow & Xu (2004) test the frequency factor in the GLA. Their participants are Mandarin speaking ESL students who are acquiring coda obstruents. In the GLA, constraint ranking is determined by frequency of overt structures in the input. The authors argue that the factor that determines IL constraint ranking is perceived frequency of overt L2 structures, not simply the absolute frequency (Broselow & Xu 2004: 138).
That is, both actual and perceived frequencies play a role in determining acquisition (Broselow & Xu 2004: 152). This study provides positive evidence for the frequency factor in the GLA and for its ability to account for L2 acquisition. The study also includes the role of perception in language acquisition.

Boersma & Escudero (2004) study the acquisition of Spanish by Dutch speakers and focus on the acquisition of an L2 vowel inventory that is smaller than that of the L1. The authors remark that learning to perceive Spanish’s five vowel system proves difficult for Dutch speakers who have to map the five vowels to discrete representations based on their L1 vowel system (Boersma & Escudero 2004: 2). This difference in vowel inventory may also pose problems for the English learner because English, like Dutch, has a larger vowel inventory than Spanish. Hence, English speakers may also experience difficulty in the perception of discrete vowel representations. Perception is based on perceived similarity, which is language specific; therefore, acquiring perceptual boundaries for the representations becomes difficult (Boersma & Escudero 2004: 3).

Comprehension begins with an initial transfer of L1 perception into the IL (Boersma & Escudero 2004: 4). The constraint ranking determines what perceived values are mapped to which vowel categories. This mapping is expected to be variable and is shown by ranking the constraints along a continuous scale. The constraints in this study are modeled according to vowel duration and formant frequencies. (Boersma & Escudero 2004: 8). The authors model real learning with a computer simulated learner in order to test the learning algorithm. They find that the initial ranking is that of the L1 grammar. The simulated learner is exposed to 10,000 Spanish vowel tokens in a year and is considered to have high motivation. The learner’s proficiency does improve, but it never
reaches nativelike L2 competence (Boersma & Escudero 2004: 15). The simulated learner resembles actual human learning, but it seems like a best case scenario. The authors conclude that this study demonstrates the link between phonological theory and computational models of language acquisition (Boersma & Escudero 2004: 21).

These studies support the use of OT in IL phonology research. OT is able to account for IL forms that are not present in the L1 or the L2 through the emergence of the unmarked. These studies have posited modifications to current OT learning models that better accommodate SLA, such as through the re-ranking of both markedness and faithfulness constraints. The studies agree that learners start off at an L1 initial state that develops towards the L2, and that during this development learners incorporate universally unmarked forms into the IL. OT models are able to allow for a more inclusive and variable model of IL phonology.

3.5.4 Summary

This section will summarize the role of the GLA in SLA research and offer some further considerations on the topic. One advantage of the GLA is that it allows for learner variation and for optionality in surface forms. Allowing for variation is an important consideration for L2 phonology, as learners experience variation along the interlanguage continuum during different stages of acquisition. In order to explain this variation it is necessary to provide for flexibility in constraint rankings. The lack of strict dominance in the GLA allows the L2 learner opportunities to observe overt forms, compare those with surface forms, and rearrange constraint rankings towards convergence.
This process is not very different from that of L1 acquisition. However, the source of overt structures is different. L2 learners are not always exposed to overt structures provided by native speakers of the target language. Instead, overt structures may be provided by an instructor who is not a native speaker of the L2 and who may or may not have reached convergence in his/her L2 speech. Also, in a communicative classroom context, overt structures are provided by other L2 learners. Hence it seems that constraint rankings would be influenced by these different types of overt forms.

Another way in which SLA differs from L1 acquisition is the presence of language transfer effects or L1 interference. The GLA is able to account for transfer in that the initial constraint ranking would be that of the L1. Although the GLA may still have to account for variable learning contexts, it has proven to be a reliable model for L2 acquisition that can explain IL processes of production and perception through a continuum of constraint rankings that allows for variation.

OT learnability models differ from contrastive models of acquisition in that there is not a one-way effect of the L1 on the L2. Rather, in OT models, L1 and L2 constraints interact and are continuously ranked and re-ranked until convergence occurs. OT models also take into account the ability to access linguistic knowledge not readily available in the L1 or the L2 through the emergence of markedness constraints in the IL that do not normally surface or are low ranked in the L1 and the L2.

3.6 Predictions

The data presented above, along with the review on L2 learnability in OT, allows for predictions to be made as to the possible constraint interactions expected in the IL of English speakers acquiring Spanish diphthongs. Moraicity and sonority are the central
explanatory principles in the analyses reviewed in this chapter. Language specific restrictions on syllable constituents and permissible segment sequences will also play a significant role in the predictions made here.

Predictions based on sonority should not be controversial. Because Spanish and English both have constraint rankings for the role of sonority in syllabification based on the universal SSG, it would be expected that L2 learners would maintain the universal sonority hierarchy in their IL. Sonority’s role in diphthongization is best seen in Colina’s (2009) mora assignment hierarchy and in Baertsch’s (2008) peak and margin hierarchies. Colina (2009) uses this hierarchy to determine which vowel is most likely to lose its mora in diphthongization. The higher the vowel, the more likely it is to lose its mora when adjacent to another vowel. Baertsch (2008) uses her peak and margin hierarchies to explain the differential behavior of CuV and CiV sequences. Her parsings result from the desire to have a more sonorous vowel in peak position. Based on the importance of sonority in these two analyses, it is expected that sonority will play a role in determining which segment L2 learners will chose to glide and if they will maintain the preferred sonority contour by forming a diphthong instead of creating an onsetless syllable through hiatus. It is expected that there will be a desire to maintain ONSET as a high-ranking constraint, as it is highly ranked in both languages. However, if hiatus is maintained, then FAITH, such as MAX-IOμ, must be more highly ranked than ONSET. ONSET is highly-ranked to achieve an optimal CV syllable.

While sonority plays an essential role in determining which segments belong to syllabic constituents, language specific restrictions also dictate allowable sequences. Diphthongization in Spanish results in C+glide sequences that are illicit in English and
that would be eliminated by the constraint rankings in Baertsch’s margin hierarchies. As predicted by the contrastive models discussed in Chapter 2, it would be expected that English speakers would have difficulty with word-initial coronal + [j] clusters that result from the gliding of /i/ in Spanish. In order for L2 learners to acquire C+glide sequences that are illicit in English, they will have to demote constraints that bar such sequences in favor of enforcing ONSET.

Another strategy for hiatus resolution that is available to the L2 learner is deletion. Low vowel deletion is attested in Chicano Spanish to avoid certain types of diphthongs that are illicit in that dialect (Colina 2009; Baković 2006). If deletion is selected as a strategy for hiatus resolution, then MAX-V will have to be demoted below ONSET.

As discussed in Chapter 2, not all vowel sequences containing a [+hi] vowel in Spanish result in a diphthong. Exceptional hiatus is attested in Spanish, in which gliding does not take place and each vowel forms its own nucleus. This pronunciation strategy should also be available to L2 learners. A tendency to split diphthongs in L2 Spanish could reflect a preference for only those diphthongs that are allowed in English, which has fewer possible diphthongs than in Spanish. By splitting a diphthong into a hiatus, learners satisfy faithfulness constraints. MAX-IOω is not violated, as the vowel keeps its mora. DEP is also satisfied over ONSET, as a new segment is not introduced between the vowels to create an onset. In order to achieve diphthongization, L2 learners will have to demote faithfulness constraints, as in Swanson (2001). It is anticipated that learners will map Spanish diphthongs onto already established categories for an acceptable nucleus, as
in Boersma & Escudero (2004) discussed above. The constraints that establish these categories would have to be ranked and re-ranked until the L2 category is stable.

There may also be other methods for resolving hiatus not used in either language or found only in certain dialects that highlight a preference for the unmarked. Because hiatus is cross-linguistically marked, it would be expected that learners would want to dissolve a sequence in hiatus, but it is possible that diphthongization is not the preferred process across all levels of the IL.

3.7 Conclusion

This chapter has reviewed both Spanish and English syllable structure from an OT perspective. Universal constraints interact in a language specific way to determine the optimal output. Constraints concerning moraicity and sonority are essential for explaining the syllabic processes in the analyses reviewed above. The analysis of dialect variation in Spanish demonstrates OT’s ability to account for variation in the output through the re-ranking of constraints. It is also evident from accounts of English syllabification that there is still much to debate that is outside of the scope of this project and not discussed here. These debates reflect pre-OT concerns that have still not been resolved within this framework.

This chapter also reviewed the Gradual Learning Algorithm as a viable model for second language acquisition. The GLA, as used in SLA, takes the L1 as the initial state and re-ranks constraints based on frequency of input to finally reach convergence in the L2. This chapter also made predictions for the acquisition of Spanish diphthongs by English speakers based on the constraint rankings proposed for Spanish and English and based on the literature concerning OT learnability models for SLA.
Chapter 4: Methodology

4.1 Introduction

This chapter outlines the research methodology for this study. Section 4.2 describes the participants. Section 4.3 describes the experimental tasks that the participants completed. The data analysis procedures are described in Section 4.4.

4.2 Participants

66 people participated in this research, 64 of whom provide the data analyzed in this study. The participants are divided into four groups. Three of the groups are comprised of current and former students at the University at Albany who were enrolled in and/or teaching Spanish classes at UAlbany at the time of the study. All participants in the first three groups are native speakers of American English. Native speakers of Spanish comprise the fourth group. Each group completed a set of tasks to be described in Section 4.3. All participants completed a background questionnaire to gather information about their experiences with Spanish. The first three groups completed the same questionnaire. The group comprised of native speakers of Spanish completed a separate questionnaire. A description of each group based on the responses to the background questionnaire appears below.

4.2.1 Beginner Group

16 students at the University at Albany participated in the Beginner Group. These students were enrolled in ASPN 101: Elementary Spanish 2 at the time of the study. Based on their course placement, these participants were considered to be at the beginning level of their acquisition of L2 Spanish. These participants ranged in age from 18-27. Seven participants reported having studied a language other than Spanish for
periods varying from four weeks to five years, the most common languages studied being French and Italian. Four participants reported having traveled to a Spanish speaking country for periods ranging from four days to two months, but they reported speaking little to no Spanish on these trips. Aside from in the classroom, participants reported hearing Spanish at work, on TV, in other classes, at a friend’s house, and while shopping. One participant reported that Spanish is sometimes spoken at home, and another reported that an uncle speaks Spanish at home. Although 16 people completed the tasks, one participant was eliminated from the study because American English was not his native dialect. This information is summarized in Fig. 4.1.

4.2.2 Intermediate Group

22 students at the University at Albany participated in the Intermediate Group. These students were enrolled in ASPN 104: Intermediate Spanish 2 or ASPN 206y: Intermediate Conversation and Oral Grammar at the time of the study. Based on their course placement, these students were considered to be at the intermediate level of their acquisition of L2 Spanish. Participants in this group ranged in age from 18-69. 11 participants had studied a language other than Spanish for periods ranging from one to eight years, French and Latin being the most commonly studied languages. 10 participants reported having traveled to a Spanish speaking country for periods ranging from one day to four months. Some participants spoke no Spanish during their travels, while others spoke Spanish to varying degrees while abroad. While the home language is English for all participants, some participants attested to other languages being spoken in their home at some time. Some participants hear some Spanish spoken at home or on the phone. Aside from in the classroom, participants reported hearing Spanish at work, at
friends’ houses, in the dorms, and at their grandparents’ house. One participant reported hearing Spanish “everywhere”. Although 22 people completed the tasks, one participant was eliminated from the study because of technical problems that resulted in an incomplete recording of the participant’s data. This information is summarized in Fig. 4.1.

4.2.3 Advanced Group

14 people participated in the Advanced Group. These participants were either undergraduate students enrolled in a 400 level Spanish course, graduate students, or Spanish instructors at the University at Albany. Based on their course enrollment or their employment as an instructor of Spanish, these participants were considered to be at the advanced level of their acquisition of Spanish. Participants in this group ranged in age from 24-34. 12 participants reported having studied a language other than Spanish for periods ranging from one semester to eight years, French and Portuguese being the most commonly studied languages. All participants in this group have traveled to a Spanish speaking country for periods ranging from one week to a total of more than two years. All participants reported speaking Spanish during their travels. Two participants reported speaking Spanish at home, but minimally or as a secondary language. Aside from the classroom, participants reported hearing Spanish at restaurants, bowling alleys, friends’ houses, on TV, on the radio, on the Internet, and at work. This information appears below in Fig. 4.1.
Fig. 4.1 Background information for Beginner, Intermediate, and Advanced Groups

<table>
<thead>
<tr>
<th></th>
<th>Beginner</th>
<th>Intermediate</th>
<th>Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants</td>
<td>14</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>Age</td>
<td>18-27</td>
<td>18-69</td>
<td>24-34</td>
</tr>
<tr>
<td>Studied other languages</td>
<td>7 out of 14 Latin, Greek, French*, Chinese, German, Italian*</td>
<td>11 out of 21 French*, Latin*, Greek, Dutch, German, Chinese, Italian, Arabic</td>
<td>12 out of 14 French*, German, Portuguese*, Italian, Latin, Catalan, and Hebrew</td>
</tr>
<tr>
<td>Travel to Spanish speaking country</td>
<td>4 out of 14 Nicaragua, Spain, Panama, Mexico</td>
<td>10 out of 21 Mexico*, Puerto Rico, Spain*, Costa Rica, Guatemala, Honduras</td>
<td>14 out of 14 Spain*, Mexico, Honduras, Colombia, Ecuador, Argentina, Chile, Peru, Costa Rica, Nicaragua, Guatemala</td>
</tr>
<tr>
<td>Places Spanish is heard</td>
<td>work*, TV, other classes, friend’s house, shopping, home</td>
<td>home, work, friends’ houses, dorms, grandparents’ house, “everywhere”</td>
<td>home, restaurants, bowling alleys, friends’ houses, TV, radio, Internet, work*</td>
</tr>
<tr>
<td>Home languages (besides English)</td>
<td>Spanish, Arabic, Jamaican</td>
<td>Spanish, Italian, Hula, French</td>
<td>Spanish</td>
</tr>
</tbody>
</table>

*most common response

4.2.4 Native Speaker Control Group

14 people participated in the Native Speaker Control Group. These participants were personal acquaintances of the investigator. All participants in this group were Spanish-English bilinguals. Participants in this group ranged in age from 20-60. Seven participants claim Colombia as their country of origin. Three participants were born in the United States, two with Puerto Rican heritage and one with Uruguayan heritage. Three participants were born in the Dominican Republic, and one was born in Cuba. The majority of participants in this group completed their primary, secondary, and undergraduate education outside of the United States. 13 participants had completed or were enrolled in graduate programs at the time of the study. 12 of these participants had
completed or were enrolled in US graduate programs. Eight participants had studied languages other than English, French and German being the most commonly studied languages. The majority of participants reported Spanish as their home language. This information appears below in Fig. 4.2.

Fig. 4.2 Background information for Native Speaker Control Group

<table>
<thead>
<tr>
<th>Number of participants</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>20-60</td>
</tr>
<tr>
<td>Country of origin</td>
<td>7 – Colombia, 3 – US, 3 – DR, 1 – Cuba</td>
</tr>
<tr>
<td>Studied other languages</td>
<td>8 out of 14 -- French*, German*, Japanese, Portuguese, Russian</td>
</tr>
<tr>
<td>Primary school</td>
<td>1 – US, 1 – Half US, 12 – outside US</td>
</tr>
<tr>
<td>Home language</td>
<td>1 – English, 3 – English/Spanish, 10 --Spanish</td>
</tr>
</tbody>
</table>

*most common response

4.3 Tasks

This section describes the two tasks completed by participants in this study. Each task consisted of reading a word, either in isolation or as part of a sentence, that contained a sequence of two vowels. Semi-spontaneous speech also forms part of the data. The participants were audio recorded using a digital voice recorder. The tasks are described below.

4.3.1 Task 1-- Nonsense words

Nonsense words have long been used to test for phonological awareness in children and adults (Bybee & Pardo 1981; Treiman 1986, 1994; Treiman & Zukowski 1988; Treiman, Goswami & Bruck 1990; Coleman & Pierrehumbert 1997; Ohala 1999; Berent, Steriade, Lennertz & Vaknin 2007; Daland, Hayes, White & Garellek 2011), but
they have not been used to study the diphthong-hiatus contrast. In this task, participants were presented with a series of 25 cards, each containing a nonsense word. All words contained a sequence of two vowels. The frequency of vowel sequences in the test words was based on the frequency of their occurrence in the 500 most frequent words in Spanish as reported in the *Frequency Dictionary of Spanish Words* (Juillard & Chang-Rodriguez 1964). The list of nonsense words consisted of 20 words with sequences containing at least one high vowel. The other five words contained a sequence of non-high vowels that would be pronounced as hiatus in Spanish. Of the expected diphthong sequences, six were rising in sonority, and three were falling in sonority. There is one of each sequence of the high vocoids /iu/ and /ui/. The expected hiatus sequences represented were /ea/, /ae/, /ao/, and the orthographically marked sequence <ía>. Not all possible diphthong and hiatus sequences were represented on this list. The goal was to include the most common diphthongs in order to more closely replicate the sequences that a beginner would most likely encounter. For example, the sequence /eu/ was not represented among the diphthongs. /eu/ is an infrequent diphthong in Spanish, and would not likely be encountered in a lower level Spanish class. The distribution of vowel sequences in the frequency dictionary and in the task stimuli can be seen in Figs. 4.5 and 4.6.

Participants were told that because the words do not currently exist in Spanish, there was no right or wrong way to pronounce them. The participants were asked to pronounce each word as well as they could, based on their knowledge of Spanish. Participants could take as much time as needed to read and pronounce each word, and multiple attempts at pronunciation were allowed. Only the final attempt was analyzed.
This task elicited conscious and controlled speech in which the participant was focused on their pronunciation. The purpose of the task was to test for knowledge of the gliding rule in Spanish, as participants could not rely on how they may have previously heard the work pronounced. The task instructions and stimuli appear below.

**Fig. 4.3 Task 1 – Nonsense words / Palabras sin sentido**

You will be presented with a series of cards, each containing a nonsense word. These words do not currently exist in Spanish, so there is no right or wrong way to pronounce these words. Please pronounce each word as well as you can based on your knowledge of Spanish.

Aquí se presenta una serie de tarjetas, cada una con una palabra que no existe en español. No hay una manera equivocada de pronunciar la palabra. Por favor, pronuncie cada palabra lo mejor que pueda a base de su conocimiento de español.

- pueve
- brazia
- agaico
- tuepa
- cuapa
- tapea
- tacua
- peago
- prafia
- biero
- teico
- lubies
- goiro
- tiaro
- paoto
- popue
- biopo
- pieto
- daeta
- gozio
- piuto
- tuisa
- buota
- tieco
- botio

**4.3.2 Task 2 – Question & Answer**

Participants were presented with a series of 15 cards, each containing a yes/no question in Spanish. Each question contained at least one word with a sequence of two vowels. As in Task 1, the frequency of vowel sequences in the test questions was based on the frequency of their occurrence in the 500 most frequent words in Spanish as reported in the *Frequency Dictionary of Spanish Words* (Juillard & Chang-Rodriguez 1964). There were a total of 24 words that contained the vowel sequences under
investigation, with a total of 25 vowel sequences. The word *biología* ‘biology’ contains two vowel sequences. The vowel sequences in the questions were the same as those of the nonsense words. The only difference was that /uo/ from the nonsense words was traded for /au/ in the question task, as /au/ is a more frequent sequence in Spanish. In both tasks, the goal was to try to include as many sequences as possible while also accounting for the frequency of the sequences in the language. The distribution of vowel sequences in the frequency dictionary and in the task stimuli appears below in Figs. 4.5 and 4.6.

Participants were asked to read the question aloud, thereby eliciting controlled and conscious speech of real words in context. After reading the question aloud, participants were asked to respond to the question in a complete sentence in Spanish. This response elicited semi-spontaneous speech. Participants were encouraged to use the words in the question to guide their response. In the case that the participant did not read the question aloud or failed to answer with a complete sentence, he/she was given a reminder of the instructions and was asked to repeat that task item. If the participant did not understand the question, a translation was provided. The task instructions and stimuli are shown below with the words under analysis highlighted in bold and the vowel sequences underlined. Note that the <h> in *ahora* ‘now’ is silent, producing the hiatus [a.o].
Fig. 4.4 Task 2 – Question and Answer / Pregunta y respuesta

You will be presented with a series of cards, each containing a question in Spanish. Please read the question aloud, and then answer the question in a complete sentence in Spanish. Feel free to use the words in the question to guide your response.

Aquí se presenta una serie de tarjetas, cada una con una pregunta en español. Por favor, lea la pregunta en voz alta, y luego responda a la pregunta en una oración completa en español.

1) ¿Duermes en clase a veces?
2) ¿Vas a ir a una fiesta hoy?
3) ¿Te gusta estudiar en la biblioteca?
4) ¿Te levantas a las seis?
5) ¿Bebes mucha agua?
6) ¿Juegas muy bien al tenis?
7) ¿Cuántas personas hay en tu familia?
8) ¿Te gusta ir al teatro?
9) ¿Vas en autobús a la universidad?
10) ¿Llevas un suéter en el invierno?
11) ¿Siempre traes los libros a clase?
12) ¿Te gusta la biología?
13) ¿Eres una persona creativa?
14) ¿Vives en una ciudad ahora?
15) ¿Necesitas llevar una chaqueta en junio?
### Fig. 4.5 Count of vocalic sequences 500 most frequent words in Juilland & Chang-Rodriguez (1964)

<table>
<thead>
<tr>
<th>Diphthong (Rising)</th>
<th>ie</th>
<th>io</th>
<th>ue</th>
<th>ia</th>
<th>ua</th>
<th>ui</th>
<th>iu</th>
<th>uo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>26</td>
<td>26</td>
<td>18</td>
<td>13</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diphthong (Falling)</th>
<th>ei</th>
<th>au</th>
<th>ai</th>
<th>oi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hiatus</th>
<th>ea</th>
<th>ae</th>
<th>ao</th>
<th>ee</th>
<th>eo</th>
<th>oe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hiatus (orthographic)</th>
<th>ía</th>
<th>aí</th>
<th>of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

### Fig. 4.6 Number of vowel sequences in Nonsense Word and Question/Answer Tasks based on frequency in Juilland & Chang-Rodriguez (1964)

<table>
<thead>
<tr>
<th>Diphthong (Rising)</th>
<th>ie</th>
<th>ue</th>
<th>io</th>
<th>ua</th>
<th>ia</th>
<th>iu</th>
<th>ui</th>
<th>uo (NW only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diphthong (Falling)</th>
<th>ei</th>
<th>ai</th>
<th>oi</th>
<th>au (QA only)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hiatus</th>
<th>ea</th>
<th>ae</th>
<th>ao</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hiatus (Orthographic)</th>
<th>ía</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>1</td>
</tr>
</tbody>
</table>

### 4.4 Phonetic analysis

The data was analyzed using Praat, version 5.3.20 (Boersma & Weenink 2012). The vowel sequences were isolated and their duration was measured in milliseconds based on the formants visible in the spectrogram. F1, F2, and F3 were measured in Hertz at each of the following points of the vowel sequence: 0.0, 0.25, 0.5, 0.75, 1.0. The
decision to take measurements as these five points was based on Smith, Flores & Gradoville (2008) in which they take measurements at 25%, 50%, and 75% points of the duration of the vowel sequence. The decision was also seen as a compromise between methodologies that take measurements at three points – beginning, transition, end – (Lehiste & Peterson 1961; Gay 1970; Borzone de Manrique 1979) and Aguilar (1999) in which measurements are taken continuously at intervals of 10 msec. Taking formant measurements at percentage points along the vowel sequence allows for the measurements to be normalized in relation to duration. The measurements were taken using a Praat script (Lennes 2003; Todt 2009; Styler 2011). The F1 and F2 measurements were sanity checked in Excel to ensure that they fell within an acceptable range for possible vowels in Spanish and English. The range for F1 was 167-1127 Hz. The range for F2 was 500-3224 Hz. The ranges were originally based on those found in Quilis & Esgueva (1983), Quilis (1993), Peterson & Barney (1952), Bradlow (1995), Hillenbrand, Gerry, Clark & Wheeler (1995), and Stevens (2000). The ranges were then adjusted to account for individual speaker differences.

4.4.1 Slope of F2

Standard deviation and the slope of F2 were tested for their ability to differentiate diphthong from hiatus. F2 was chosen as the most useful formant due to its wide range, approximately 500-3000 Hz, and its use in previous analyses (Gay 1970; Borzone de Manrique 1976, 1979; Aguilar 1999; Colantoni & Limanni 2010). The decision to test for standard deviation and slope, which both measure change from a given point, was based

---

8 The script used in this research is a combination of the scripts mentioned in the citation along with modifications to suit this study’s data set. Thank you to Janet Smith for sharing Mietta Lennes’ script containing Kirsten Todt’s modifications.
on previous research. Gay (1970) studied English diphthongs and questioned if the F2 transition is the primary cue for identifying diphthongs versus single vowels. He found that the F2 rate of change is a distinguishing factor (Gay 1970: 87). Borzone de Manrique (1976, 1979) also used rate of change for investigating Spanish diphthongs. Although F2 rate of change is useful for distinguishing different types of diphthongs, it is not the sole element in distinguishing between diphthong and hiatus. In fact, Borzone de Manrique found that “the rate of change […] showed neither a significant nor a consistent difference between the vowel sequences which form a diphthong and those which do not” (1979: 203). Aguilar (1999) relied on the curvature of the formant trajectory in her study of Spanish diphthongs. She found that F2 trajectory is a reliable indicator of the diphthong-hiatus contrast, but it is more convincingly used in confirming vowel reduction in diphthongs, as reduced diphthongs have a flatter trajectory (Aguilar 1999: 73).

Colantoni & Limanni (2010) studied the diphthong-hiatus contrast in Argentinian Spanish in a contact situation. They found that the slope of F2 was a poor indicator of the diphthong-hiatus divide in native speaker perception (Colantoni & Limanni 2010: 31). The present study agrees with more recent research (Colantoni & Limanni 2010; MacLeod 2012) and with work by Borzone de Manrique (1979) that slope does not reveal significant findings.

Nevertheless, the slope of F2 was useful for this research in determining which sequences were monophthongs resulting from elision or coalescence. The slope of F2 was calculated between each of the points mentioned above (0.0, 0.25, 0.5, 0.75, 1.0), and the highest slope was isolated. If the highest slope was low, between 1.5-5.0 Hz/ms depending on the sequence, that sequence was counted as a monophthong. This
categorization was confirmed by plotting the F2 values in Excel. The particular monophthong used by the speaker was determined by the formant frequencies.

4.4.2 Duration

Duration was used to differentiate the remaining sequences into diphthongs and hiatus. Duration is the most cited indicator in making a diphthong-hiatus distinction, and the literature confirms that vowel sequences in hiatus have a longer duration than those pronounced as a diphthong (Lehiste & Peterson 1961; Gay 1970; Borzone de Manrique 1979; Aguilar 1999; Hualde & Prieto 2002; Face & Alvord 2004; Garrido 2007; Smith, Flores & Gradoville 2008; Colantoni & Limanni 2010; MacLeod 2012). Duration of the individual components of the vowel sequence is frequently measured along with the duration of the entire sequence to decipher the role of the vowel target or the length of the transition (Lehiste & Peterson 1961; Gay 1970; Borzone de Manrique 1976, 1979; Colantoni & Limanni 2010; MacLeod 2012). Duration also plays a role in studies mentioned above that utilize rate of change, as slope is a function of duration. In sum, duration is almost always present in an acoustic account of diphthongization.

While duration has been established as a useful indicator in the diphthong-hiatus contrast, it is necessary to define its parameters for defining the contrast in Spanish. Hualde & Prieto were unable to establish a durational threshold at which to mark off the distinction (2002: 233). Face & Alvord replicate Hualde & Prieto (2002) and find that there is overlap in diphthong and hiatus durations due to the phonetic context and individual speaker differences (2004: 558). There is also the necessity to control for vowel quality when measuring duration, as different sequences have different durations (Hualde & Prieto 2002: 222). Although thresholds are not available in the current
literature, there are comprehensive accounts of average durations for diphthong and hiatus sequences. Borzone de Manrique (1979) presents duration times for 14 diphthongs in Spanish according to speech rate in two different tasks, reading isolated words and reading words in a sentence. Aguilar (1999) presents the average duration times for the diphthong and hiatus pronunciations of vocalic sequences of rising sonority according to the task types used in her research, which are a reading task and a dialogue task. As for sequences of non-high vowels, Garrido (2007) records average duration times for the sequence /eo/ in a reading and a narration task. Smith, Flores & Gradoville (2008) present average duration times for the sequences /ae/, /oa/, /ea/, and /oe/ in a narration task.

These average duration times are used in the present methodology as thresholds for establishing a diphthong-hiatus contrast. There are different duration times depending on task and rate of speech. The data resulting from the Question & Answer task had slightly different mean durations in both the question and answer portions of the task, so the two types of data resulting from this task will be analyzed using slightly different durational thresholds. To account for the data in the Question & Answer task, the moderate speech average resulting from the isolated word task in Borzone de Manrique (1979) and the average diphthong and hiatus durations from the reading corpus in Aguilar (1999) were averaged for each sequence of rising sonority. For the sequences of falling sonority and the sequences containing only high vowels, the durations from the isolated word task in Borzone de Manrique (1979) are used as they are the only data available. The Question data had longer duration, so the threshold had to account for this difference. The Answer averages were multiplied by a factor of 1.03 to account for the slightly
longer average duration for the Question data. The equation used to arrive at this factor is described in more detail below. Similarly, the durations for the sequences /oi/ and /ai/ were normalized for duration by a factor of 1.2 and 1.4, respectively. It was found that these two sequences were longer than the other sequences because of their position at the end of a sentence or because the words tended to be drawn out as participants were thinking of the next word in their response.

For the sequences of two non-high vowels, the durations from Smith, Flores & Gradoville (2008) are used. However, these durations result from a narration task, which would have shorter durations than a reading task. The type of data resulting from the Question and Answer task more closely resembles that of a reading task. In order to normalize duration according to task, an equation is used based on the methodology for normalizing duration among speakers in MacLeod (2012). To consider what a likely difference between a reading and a narration task would be, the average durations for all sequences in the two tasks in Aguilar (1999) were compared. The average duration for all the data in the dialogue task was 139 ms, and the average duration for the reading task was 167 ms. Using the following equation, it was possible to find a factor by which to account for the increase in duration that a reading task would impose. The equation is: 1 + ((167-139)/139) = 1.2. The duration values from Smith, Flores & Gradoville (2008) were multiplied by a factor of 1.2 to account for the durational difference in task type. The reading task in Garrido (2007) provided the duration for the sequence /eo/.

The duration for <ía> was the result of averaging the data for /ía/ in hiatus in Aguilar (1999), Hualde & Prieto (2002), and Face & Alvord (2004). Although these three studies are examining the /ía/ hiatus in which the /a/ is stressed, it is the available
data that most closely resembles the data here. It may seem that the durational thresholds have been arrived at in a rather piecemeal fashion, but at this time, there is no comprehensive account of the possible pronunciations of Spanish vowel sequences. The durational thresholds and the F2 slope thresholds for the Question & Answer task are shown in the figure below.

**Fig. 4.7** Average duration and slope used to determine diphthong-hiatus contrast: Question data

<table>
<thead>
<tr>
<th>Vocoid Sequence</th>
<th>ai</th>
<th>ei</th>
<th>oi</th>
<th>au</th>
<th>ia</th>
<th>ie</th>
<th>io</th>
<th>ua</th>
<th>ue</th>
<th>uo</th>
<th>iu</th>
<th>ui</th>
<th>eo</th>
<th>ea</th>
<th>ae</th>
<th>ao</th>
<th>oe</th>
<th>fa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration (ms)</td>
<td>223</td>
<td>162</td>
<td>216</td>
<td>161</td>
<td>179</td>
<td>135</td>
<td>168</td>
<td>146</td>
<td>178</td>
<td>137</td>
<td>185</td>
<td>136</td>
<td>159</td>
<td>122</td>
<td>199</td>
<td>163</td>
<td>NA</td>
<td>181</td>
</tr>
<tr>
<td>Slope F2 (Hz/ms)</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3.5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1.5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>NA</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 4.8** Average duration and slope used to determine diphthong-hiatus contrast: Answer data

<table>
<thead>
<tr>
<th>Vocoid Sequence</th>
<th>ai</th>
<th>ei</th>
<th>oi</th>
<th>au</th>
<th>ia</th>
<th>ie</th>
<th>io</th>
<th>ua</th>
<th>ue</th>
<th>uo</th>
<th>iu</th>
<th>ui</th>
<th>eo</th>
<th>ea</th>
<th>ae</th>
<th>ao</th>
<th>oe</th>
<th>fa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration (ms)</td>
<td>216</td>
<td>157</td>
<td>210</td>
<td>156</td>
<td>174</td>
<td>131</td>
<td>163</td>
<td>142</td>
<td>173</td>
<td>133</td>
<td>180</td>
<td>132</td>
<td>154</td>
<td>118</td>
<td>193</td>
<td>158</td>
<td>163</td>
<td>176</td>
</tr>
<tr>
<td>Slope F2 (Hz/ms)</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3.5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1.5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>NA</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

The average duration for the sequences in the Nonsense Word task was 35% longer than average duration in the Question and Answer task. The same duration values used in the Question and Answer task were used for the Nonsense Words, but multiplied by a factor of 1.35 to account for the decrease in speech rate. The durational thresholds appear below in Fig. 4.9.

**Fig. 4.9** Average duration and slope used to determine diphthong-hiatus contrast: Nonsense Words task

<table>
<thead>
<tr>
<th>Vocoid Sequence</th>
<th>ai</th>
<th>ei</th>
<th>oi</th>
<th>au</th>
<th>ia</th>
<th>ie</th>
<th>io</th>
<th>ua</th>
<th>ue</th>
<th>uo</th>
<th>iu</th>
<th>ui</th>
<th>eo</th>
<th>ea</th>
<th>ae</th>
<th>ao</th>
<th>oe</th>
<th>fa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration (ms)</td>
<td>208</td>
<td>212</td>
<td>230</td>
<td>211</td>
<td>235</td>
<td>177</td>
<td>220</td>
<td>192</td>
<td>234</td>
<td>180</td>
<td>243</td>
<td>178</td>
<td>208</td>
<td>159</td>
<td>261</td>
<td>213</td>
<td>220</td>
<td>238</td>
</tr>
<tr>
<td>Slope F2 (Hz/ms)</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3.5</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1.5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>NA</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
The thresholds for each sequence were entered into Excel and the data was categorized according to the thresholds in Fig 4.7, 4.8, and 4.9 as being a monophthong, diphthong, or hiatus. Sequences with a slope value falling below the slope threshold were counted as monophthongs. If the duration fell below the durational threshold, the sequence was counted as a diphthong. If the duration fell above the durational threshold, the sequence was counted as a hiatus. While averaging the diphthong and hiatus duration times creates a gray area for distinguishing between the two pronunciations, some overlap is to be expected based on previous research (Hualde & Prieto 2002; Face & Alvord 2004). While there may be some diphthongs that land in the hiatus category and some hiatus that land in the diphthong category, using average thresholds allows for establishing trends in the data as to the distribution of diphthong versus hiatus pronunciation for a given vowel sequence.

4.5 Conclusion

64 people participated in this research and completed a Nonsense Word and a Question & Answer task. Slope of F2 and duration of the vowel sequences indicated which pronunciation was used – monophthong, diphthong, or hiatus. This methodology was used to establish trends in the data for subsequent theoretical analysis. Duration was the principle variable used in this research, and its use for diphthong-hiatus analysis has been a mainstay in the research on vowel sequences. However, the use of duration is not without qualification in the literature. Borzone de Manrique maintains that the “interpretation of a vowel sequence […] seems to be a function of the totality of the speaker’s linguistic system” (1979: 205). Hualde & Prieto acknowledge that duration and formant values are necessary to make a full account of the distinction (2002: 233).
Face & Alvord call for the inclusion of lexical frequency effects to improve on the explanations offered by phonetic analyses (2004: 563-564). Smith, Flores & Gradoville note the general “lack of an objective method with which to classify vowel combinations” and to the “lack of a fine-grained analysis of the acoustic correlates of tongue movements involved in the production of vowels and their transitions” (2008: 2, 4). They also note that “criteria are not yet in place to differentiate hiatus from diphthong using formant measurements alone” (Smith, Flores & Gradoville 2008: 7). It appears that there is no best way to differentiate between diphthong and hiatus pronunciations at this point, but research has proven that duration is capable of establishing pronunciation tendencies in the diphthong-hiatus divide. Hence, in this research, duration with the support of formant values forms the basis of the phonetic analysis. The purpose was not to produce a fine-grained phonetic analysis, but rather to establish trends in the pronunciation of vowel sequences in Spanish by L2 learners. The next chapter presents the pronunciation trends resulting from the phonetic analysis.
Chapter 5: Results

5.1 Introduction

This chapter presents the data resulting from the completion of the two tasks described in Chapter 4, which were a Nonsense Words task and a Question & Answer task. Based on previous research, the anticipated pronunciations for vowel sequences of two vowels are a hiatus, a diphthong, or a monophthong resulting from coalescence or reduction (Stockwell & Bowen 1965; Quilis 1993; Hualde 2005). This chapter begins by describing the data set. After this description, the chapter presents the results from the acoustic analysis. First, the slope of F2 was used to determine which sequences were pronounced as monophthongs. Then the duration of each sequence was used to differentiate diphthong and hiatus pronunciations. Participants also produced sequences other than the anticipated monophthong-diphthong-hiatus alternations. These pronunciations are described, as well. Last, results from comparing F2 values between the three participant groups are presented.

5.2 Data Set

4902 tokens of vowel sequences were analyzed in this corpus. All of the sequences occur word-internally. The data is grouped by Nonsense Words, Question, and Answer. Although the Question & Answer task was one task, the data from the Question and Answer portions of the task will be analyzed separately. The data in each group accounts for approximately a third of the total data. The totals for each task are seen in Fig. 5.1.
Although the Nonsense Words and the Question & Answer tasks each had 25 test items, they do not have the same number of total tokens. Upon producing the nonsense words, some participants inserted additional vowels into the word, thereby increasing the total number of tokens. For example, the nonsense word ‘agaico’ was on occasion pronounced as [agaisio]. The number of tokens resulting from the Question portion of the task is lower than expected because participants sometimes omitted a word in the question or because the participant failed to repeat the question when prompted.

As described in Chapter 4, the occurrence of each sequence in the stimuli was weighted for its frequency in Spanish. The number of tokens per sequence type in the results is roughly proportional to the number of items per each sequence type in the stimuli. See Fig. 5.2 for the total number of tokens by type of sequence, categorized by a rise or a fall in sonority for sequences with a high vowel, sequences containing two high vowels, and sequences without a high vowel.

**Fig. 5.1 Total number of tokens by task**

<table>
<thead>
<tr>
<th>Task</th>
<th>Number of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonsense Words</td>
<td>1644</td>
</tr>
<tr>
<td>Question</td>
<td>1619</td>
</tr>
<tr>
<td>Answer</td>
<td>1639</td>
</tr>
<tr>
<td>TOTAL</td>
<td>4902</td>
</tr>
</tbody>
</table>

**Fig. 5.2 Total number of tokens by type of sequence**

<table>
<thead>
<tr>
<th>Type of sequence</th>
<th>Number of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rising</td>
<td>2801</td>
</tr>
<tr>
<td>Falling</td>
<td>829</td>
</tr>
<tr>
<td>iu/ui</td>
<td>387</td>
</tr>
<tr>
<td>[− high] vowels</td>
<td>885</td>
</tr>
<tr>
<td>TOTAL</td>
<td>4902</td>
</tr>
</tbody>
</table>
Fig. 5.3 presents the total number of tokens for each sequence, which is roughly proportional to the number of items of each sequence in the stimuli. The sequences of two high vowels are included in the rising group, as it is the first segment that is pronounced as a glide during diphthongization, thereby lowering its sonority (Navarro Tomás 1932: 65; Hualde 2005: 80). Rising diphthongs form the largest group of data, followed by falling diphthongs and sequences of two [- high] vowels, in that order.

### Fig. 5.3 Total number of tokens per sequence by sequence type

<table>
<thead>
<tr>
<th>Rising</th>
<th>Number of tokens</th>
<th>Falling</th>
<th>Number of tokens</th>
<th>[- high] vowels</th>
<th>Number of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>ie</td>
<td>762</td>
<td>oi</td>
<td>331</td>
<td>ea</td>
<td>371</td>
</tr>
<tr>
<td>io</td>
<td>581</td>
<td>ai</td>
<td>224</td>
<td>ao</td>
<td>150</td>
</tr>
<tr>
<td>ia</td>
<td>408</td>
<td>ei</td>
<td>180</td>
<td>ae</td>
<td>139</td>
</tr>
<tr>
<td>ue</td>
<td>620</td>
<td>au</td>
<td>94</td>
<td>eo</td>
<td>18</td>
</tr>
<tr>
<td>ua</td>
<td>364</td>
<td></td>
<td></td>
<td>oe</td>
<td>1</td>
</tr>
<tr>
<td>uo</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iu</td>
<td>205</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ui</td>
<td>182</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>3188</td>
<td>TOTAL</td>
<td>829</td>
<td>TOTAL</td>
<td>679</td>
</tr>
</tbody>
</table>

### 5.3 Slope Results

This section presents the results based on the slope of F2. The slope of F2 was used to measure the relative flatness of the curvature of F2. The slope measurement used for each sequence can be seen in Figs. 4.7, 4.8, and 4.9. If the slope of F2 was below a certain low value, which was determined by plotting the F2 values in Excel, the sequence was considered to be a monophthong. See Fig. 5.4 for an example of the plot of /ue/ pronounced as a monophthong. Both monophthong and diphthong pronunciations are shown below for comparison. The frequency of F2 was used to determine the vowel quality of the monophthong. The monophthongs resulted from elision of one of the
segments in the vowel sequence, such as in the figure below in which speakers opted for /u/ or /e/ over /ue/. Coalescence also contributed to the creation of monophthongs, such as the /o/ pronunciation below in which the two sounds are combined to form an intermediary vowel. In the case of /o/ for /ue/, the features [+ round] and [+ back] are retained from /u/, but the sound lowers to approach the [- high] value of /e/. In other instances, some sounds were reduced to schwa or to some other reduced vocalic utterance.

Fig. 5.4 Monophthong and diphthong pronunciations of /ue/

There were a total of 82 monophthongs produced in the entire corpus. Participants at the beginning level of acquisition produced the most monophthongs, followed by the intermediate level participants. Advanced L2 learners and native
speakers produced the same number of monophthongs. These results can be seen in Fig. 5.5.

**Fig. 5.5 Number of monophthongs by level of acquisition**

<table>
<thead>
<tr>
<th>Level</th>
<th>Number of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>32</td>
</tr>
<tr>
<td>I</td>
<td>28</td>
</tr>
<tr>
<td>A</td>
<td>11</td>
</tr>
<tr>
<td>N</td>
<td>11</td>
</tr>
<tr>
<td>TOTAL</td>
<td>82</td>
</tr>
</tbody>
</table>

Deletion was the most common process used to create monophthongs, the high vowel being deleted most frequently. In sequences of two high vowels, the first vowel was deleted most frequently. In sequences of two non-high vowels, the second vowel was deleted most frequently, regardless of vowel quality. This data contradicts Aguilar (1999), in which the high vocoid is maintained more frequently than the syllable nucleus. However, the primacy of the first vowel is maintained in sequences of two non-high vowels (Aguilar 1999: 71). Reduction to schwa or a lax vowel was the next most common way to monophthongize a vowel sequence. Coalescence occurred least frequently. Fig. 5.6 shows the distribution of monophthongization process.

**Fig. 5.6 Monophthongization processes**

<table>
<thead>
<tr>
<th>Process</th>
<th>Number of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td>64</td>
</tr>
<tr>
<td>Reduction</td>
<td>14</td>
</tr>
<tr>
<td>Coalescence</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>82</td>
</tr>
</tbody>
</table>
5.4 Duration Results

Duration was used to differentiate diphthongs from hiatus. Durational thresholds for each sequence were based on previous research (Borzone de Manrique 1979; Aguilar 1999; Garrido 2007; Smith, Flores & Gradoville 2008). See section 4.4.2 for details on how thresholds were set. If a sequence’s duration was above the threshold set for that particular sequence, the sequence was counted as a hiatus. If its duration fell below the threshold, it was counted as a diphthong. As was stated in section 4.4.2, the purpose was to establish trends in the data for diphthongization or hiatus. Below are the results for the whole corpus of 4902 vowel sequences based on duration.

Fig. 5.7 Results for all sequences

<table>
<thead>
<tr>
<th>Sound</th>
<th>Result</th>
<th>Task Independent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>I</td>
</tr>
<tr>
<td>ALL</td>
<td>MONOPHTHONG</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>DIPHTHONG</td>
<td>32%</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>HIATUS</td>
<td>66%</td>
<td>48%</td>
</tr>
</tbody>
</table>

Fig. 5.7 reveals that diphthongs and hiatus occurred almost equally in the data at 50% and 48%, respectively, and that monophthongs made up a small percentage of the data. The stimuli for each task contained 80% sequences that would be pronounced as diphthongs according the gliding rule in Spanish and 20% that would be pronounced as hiatus. It seems that the hiatus total is exaggerated based on the stimuli. However, upon looking at the data for each level of acquisition in comparison with the native speaker control group, it becomes evident that the high percentage of hiatus pronunciation relative to the stimuli is due to the tendency towards hiatus in the Beginner group. The tendency towards diphthongization remains strong even in the other two L2 groups. In
the Native Speaker control group, the tendency is still not especially low at over 30%.
Again, these totals seem to be exaggerated.

Sorting the data by task allows for an explanation of the high hiatus percentage, as can be seen in Fig. 5.8. The results for the Question & Answer task are similar. Participants in the Beginner group produced the most hiatus, while those in the Native Speaker Control group produced the most diphthongs. The Intermediate and Advanced groups show increasingly higher rates of diphthongization, correlating with an increase in the level of acquisition. However, neither of the groups reached the level of diphthongization of the control group. In the Question portion of the task, Beginners produced 69% of the sequences in hiatus, and Native Speakers produced 32% of the sequences as hiatus. The Intermediate and Advanced groups fall in between, and hiatus rates decrease with an increase in level of acquisition. In the Answer portion of the task, the Beginner group produced 64% of the sequences as a hiatus, while the Native Speakers produced 20% of the sequences in hiatus. The Answer results for the Native Speaker group come close to matching the stimuli (80% diphthongs, 20% hiatus), with 79% of the sequences pronounced as diphthongs and 21% as hiatus. These results agree with MacLeod (2012) in that L2 learners produce more native-like diphthongs in Spanish as their level of acquisition increases.

The results from the Nonsense Words task do not show great change between the levels of acquisition. There is also little difference between the L2 groups and the Native Speaker Control group. The range for diphthongs is between 38%-48% for all participant groups, and 50%-61% of the Nonsense Words sequences were pronounced in hiatus. In total, there are more hiatus than diphthongs. Concerning the entire corpus, the results for
the Nonsense Words task and the high percentage of hiatus in the Beginner and Intermediate groups are what accounts for the higher than expected percentage of hiatus relative to the stimuli. The results for all three tasks can be seen below in Fig. 5.8. The percentages are rounded up or down to the nearest whole number.

**Fig. 5.8 Results by task**

<table>
<thead>
<tr>
<th>Sound</th>
<th>Result</th>
<th>Question</th>
<th>Answer</th>
<th>Nonsense</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>MONOPHTHONG</td>
<td>B 3%</td>
<td>I 2%</td>
<td>A 1%</td>
</tr>
<tr>
<td></td>
<td>DIPHTHONG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HIATUS</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results in Fig. 5.8 above correspond with the results in Fig. 5.9 below, which shows the average duration for each participant group for each set of data. In both the Question and Answer data sets, the average duration decreases as the level of acquisition increases. As diphthongs have a shorter duration than hiatus, it would be expected that the average duration would go down as the percentage of diphthongization goes up.

Hence, the average duration results confirm the diphthongization data for the Question & Answer task in Fig. 5.8 above. The average duration also corresponds with diphthongization data for Nonsense Words. The diphthongization data shows little change between levels of acquisition, as does the average duration below. See the average duration below in Fig. 5.9.

**Fig. 5.9 Average duration for each group by task**

<table>
<thead>
<tr>
<th>Task</th>
<th>Question</th>
<th>Answer</th>
<th>Nonsense Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>B I A N</td>
<td>B I A N</td>
<td>B I A N</td>
</tr>
<tr>
<td>Duration (ms)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>117</td>
<td>215 172 159 141</td>
<td>218 167 155 131</td>
<td>248 221 227 230</td>
</tr>
</tbody>
</table>
5.4.1 Results by Segment

Because the Nonsense Words data departs from the patterns presented in the Question & Answer data and in the stimuli, only the Question & Answer data will be reported in this section. The sequences containing at least one high vowel that were most frequently realized as a hiatus were /ia/, /ie/, and /ei/. The sequences that were most frequently realized as a diphthong were /ue/, /iu/, and /au/. Hence, it is not possible to determine that one type of diphthong, either rising or falling, is easier to acquire. Fig. 5.10 shows the percentages of diphthongs per sequence for all speakers.

**Fig. 5.10 Percent diphthongization by segment (all speakers)**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>ia</th>
<th>ie</th>
<th>io</th>
<th>ua</th>
<th>ue</th>
<th>iu</th>
<th>ui</th>
<th>ai</th>
<th>ei</th>
<th>oi</th>
<th>au</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Diphthong</td>
<td>38%</td>
<td>49%</td>
<td>53%</td>
<td>53%</td>
<td>87%</td>
<td>86%</td>
<td>66%</td>
<td>62%</td>
<td>52%</td>
<td>60%</td>
<td>74%</td>
</tr>
</tbody>
</table>

**bold** = highest percentage of diphthongs  
**italics** = lowest percentage of diphthongs

The results are similar if only the L2 groups are considered. While the percentage of diphthongs per sequence is lower for each sequence, the relative ease of diphthongization remains the same for each segment. Fig. 5.11 shows the percentage of diphthongization per sequence for the L2 groups.

**Fig. 5.11 Percent diphthongization by segment (L2 groups)**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>ia</th>
<th>ie</th>
<th>io</th>
<th>ua</th>
<th>ue</th>
<th>iu</th>
<th>ui</th>
<th>ai</th>
<th>ei</th>
<th>oi</th>
<th>au</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Diphthong</td>
<td>30%</td>
<td>41%</td>
<td>51%</td>
<td>52%</td>
<td>84%</td>
<td>82%</td>
<td>55%</td>
<td>57%</td>
<td>50%</td>
<td>53%</td>
<td>68%</td>
</tr>
</tbody>
</table>

**bold** = highest percentage of diphthongs  
**italics** = lowest percentage of diphthongs

5.4.2 Other Realizations

Within the data accounted for there are realizations other than the anticipated monophthong, diphthong, and hiatus pronunciations. One of these unanticipated
pronunciations was metathesis of the vowel sounds. For example, *seis* ‘six’ was pronounced as [síes], *ciudad* ‘city’ as [suiða], *agaico* as [agiako], *piuto* as [puito], as a diphthong or hiatus. In some cases, a high vowel was displaced to another syllable. For example, *agaico* was pronounced as [agasio], *goiro* as [gorío], *lubies* as [luðbes]. In other cases, a vowel was inserted to create an additional vowel sequence or a triphthong. For example, *agaico* was pronounced as [agaisio], *tuisa* as [tuisia], *goiro* as [gúairo], *tacua* as [taküia], *piuto* as [pajuto], *biopo* as [bajopo], *biología* ‘biology’ as [bajoloxia].

Sometimes the vowel sequence was changed. For example, *goiro* was pronounced as [gúaro], *tacua* as [takío], *cuapa* as [küepa], *lubies* as [luíbes], *popue* as [popei], [popú] or [popui], *paoto* as [pejoto], *muy* ‘very’ as [moi]. <ju> sequences were sometimes pronounced as [ju]. For example, *juegas* ‘you play’ was pronounced as [jugas], *junio* ‘June’ as [junio]. Similarly, *universidad* ‘university’ was pronounced as [juniversidad].

In addition to vowel processes, participants sometimes inserted consonants. For example, *duermes* ‘you sleep’ was pronounced as [dorimes], *daeta* as [dalita], *piuto* as [piʔuto], *goiro* as [goʔiro], *paoto* as [paʔoto]. All of the these pronunciations occurred in the L2 groups, with the exception of metathesis which also occurred twice in the Nonsense Words in the Native Speaker Control group. See Fig. 5.12 for a summary of these pronunciations.
As explained in Chapter 4, the F1, F2, and F3 values were taken at five different points along the duration of the vowel sequence. This section considers the F2 values taken at the 25% and 75% points in the duration of each vowel sequence. F2 is useful in differentiating between vowels because of its wide range in the spectral space. In fact, Martínez Celdrán (1995: 217) found that F2 was more useful than F1 for distinguishing between vowels in Spanish. The 25% and 75% points are chosen here in order to best capture the two different vowel sounds in each sequence without interference from adjacent consonants at the beginning and end points.

The figures below present average F2 values for each participant group at the 25% and 75% points of each sequence. Because the vowels in this data are adjacent to each other, these sequences may be centralized or otherwise altered to accommodate the other vowel in the sequence. Furthermore, it is expected that vowels in diphthongs are closer in the vowel space than they are in hiatus, as vowel sequences that take place in a shorter time span are usually closer in the vowel space (Borzone de Manrique 1979; Aguilar 1999). Changes in the formant frequency from what would be expected for a steady state vowel are evident in this data.
Quilis & Esgueva (1983) and Martínez Celdrán (1995) show average formant frequencies for men and women for each vowel sound in Spanish. Peterson & Barney (1952) and Hillenbrand, Getty, Clark & Wheeler (1995) do the same for English. The average F2 values from these studies are compared with those in Fig. 5.13 in order to determine to what extent vowel quality is altered when vowels are in contact. The F2 value of /a/ raises to accommodate both /i/ and /e/, but stays within its normal range when in proximity to /o/ and /u/. Similarly, the F2 of /o/ raises to accommodate /e/ and /i/, but stays within its expected range when accompanied by /u/ and /a/. The F2 of /e/ generally lowers when in contact with /u/, and there is a slight lowering with /a/. It maintains its expected range with /i/. The F2 of /i/ lowers with /a/, /o/, and /u/. There is only a slight lowering with /e/ and when in hiatus with /a/. The F2 of /u/ raises with /a/, /e/, and /i/, increasing in frequency with each sound, respectively. It stays within its range with /o/. The data here confirms previous research that vowels in contact do move to accommodate each other’s pronunciation. However, the frequency change seemed to depend more on the quality of the adjacent vowel, than on whether the vowel was in a hiatus or diphthong.
The data in Fig. 5.13 shows that for the majority of the sequences, the Advanced group best approximates the F2 values for the Native Speaker group, while the Beginner group values are farthest away from those of the Native Speaker group. In general, the F2 values for Beginner and Intermediate groups are lower than those of the Advanced and Native Speaker groups. However, upon separating the male and female data for each sequence, no real pattern emerges to indicate an effect for level of acquisition. On average, men have lower voice frequency than women due to physiological differences. The Beginner group had twice as many men than women, and the Intermediate group had a nearly equal number of men and women. On the other hand, the Advanced and Native Speaker groups had two and three men, respectively, and the rest of the participants were women.
In only a few sequences is there any pattern to indicate an effect for level of acquisition on vowel quality. These sequences can be seen below in Fig. 5.14. The measurements showing the level of acquisition effect are shaded. For example, /a/ in the sequences /ai/ and <ia> showed changes according to level of acquisition. In both sequences, /a/ was higher for Native Speakers than for the L2 groups. The Beginners had the lowest F2 value, and that value rose toward the Native Speaker value as the level of acquisition increased. The Advanced group was able to reach nearly identical F2 values to those of the Native Speaker group for /a/ in these sequences. /u/ in the sequences /ue/ and /iu/ also showed an effect for level of acquisition. The average F2 value for the Native Speaker group was lower than for the L2 groups. Again, the Beginner group was farthest away from the Native Speaker group, with the other two groups approaching the Native Speaker value as the level of acquisition increased. The sequence /ue/ showed level of acquisition effects for both vowels. /e/ values were lower for the Native Speaker group than they were for the L2 groups, with values approaching Native Speaker values as the level of acquisition increased.

**Fig. 5.14 Sequences with level of acquisition effect**

<table>
<thead>
<tr>
<th></th>
<th>25% B</th>
<th>I</th>
<th>A</th>
<th>N</th>
<th>75% B</th>
<th>I</th>
<th>A</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>ai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1730.702</td>
<td>1733.051</td>
<td>1813.206</td>
<td>1813.849</td>
<td>Female</td>
<td>1939.772</td>
<td>2043.867</td>
<td>2109.324</td>
</tr>
<tr>
<td>Male</td>
<td>1633.475</td>
<td>1621.698</td>
<td>1878.505</td>
<td>1737.366</td>
<td>Male</td>
<td>1878.683</td>
<td>1945.375</td>
<td>1870.211</td>
</tr>
<tr>
<td>fa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2491.614</td>
<td>2553.128</td>
<td>2435.932</td>
<td>2518.71</td>
<td>Female</td>
<td>1680.652</td>
<td>1761.907</td>
<td>1807.808</td>
</tr>
<tr>
<td>Male</td>
<td>2137.236</td>
<td>2059.953</td>
<td>2173.222</td>
<td>1951.566</td>
<td>Male</td>
<td>1440.598</td>
<td>1591.068</td>
<td>1662.332</td>
</tr>
<tr>
<td>ue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1601.363</td>
<td>1434.811</td>
<td>1421.58</td>
<td>1312.155</td>
<td>Female</td>
<td>2084.644</td>
<td>1989.191</td>
<td>2027.145</td>
</tr>
<tr>
<td>Male</td>
<td>1233.291</td>
<td>1118.892</td>
<td>1055.666</td>
<td>1058.292</td>
<td>Male</td>
<td>1618.531</td>
<td>1572.956</td>
<td>1415.741</td>
</tr>
<tr>
<td>iu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2280.247</td>
<td>2369.731</td>
<td>2381.008</td>
<td>2217.872</td>
<td>Female</td>
<td>2107.494</td>
<td>1987.353</td>
<td>2001.135</td>
</tr>
</tbody>
</table>
A question to be answered in this data is whether the F2 values are closer to English or Spanish vowels in the L2 data. The average F2 values reported in Quilis & Esgueva (1983), Martínez Celdrán (1995), Peterson & Barney (1952), and Hillenbrand, Clark, Getty & Wheeler (1995) were used to determine if the vowels in the present study fall more in the English range or the Spanish range for the particular vowel sound. In addition to the studies mentioned above, Bradlow (1995) offers a comparative acoustic analysis of the Spanish and English vowel systems with tokens produced by male native speakers of both languages. She compares /i/, /e/, /o/, and /u/, which are the common vowels between the two languages. Spanish /a/ is a central vowel, whereas English /a/ is a back vowel, so Bradlow (1995) did not include /a/ and /a/ in her comparison. However, the difference in these two sounds may provide insight for the L2 pronunciation of Spanish /a/, as comparative accounts suggest (Stockwell & Bowen 1965; Whitely 2002). Because /a/ is a back vowel, it has a lower F2 value than /a/. This difference may explain the lower F2 values for Spanish /a/ in the L2 groups in general, and especially the Beginner group in the /ai/ and <ía> sequences. However, /a/ is not lower in all sequences.

Concerning the /u/ sequences, Bradlow (1995: 1918) found that English /u/ was approximately 200-250 Hz higher than Spanish /u/. The L2 groups in the /ue/ and /iu/ sequences mentioned above have higher F2 values than the Native Speaker group, ranging from nearly identical values for the Advanced group to a difference of more than 300 Hz for the Beginner group. However, this difference does not exist for all sequences containing /u/.
As is the case for /u/, Bradlow (1995: 1919) found that F2 is higher in English than in Spanish for the other three vowels that she compared. This difference is not consistent with the data in the present study. The present study examined vowels in contact that are transitioning and accommodating each other’s pronunciation. Hence, the differences observed in Bradlow’s study that occurred in CVC and CVCV sequences may not be present in the same way in CVV sequences, especially when one of those vowels is phonetically a glide. F2 of the high glide is higher than for the vowel /i/, and the low glide has a lower F2 value than for the vowel /u/ (Stevens 1998: 523, 526).

The results show that differences in F2 frequency are not explained by level of acquisition for most sequences. Rather, the differences that appear in Fig. 5.13 are a result of the distribution of the sex of the speakers in each group. /a/ and /u/ do show some level of acquisition effects, but the effects are not consistent in that they do not occur in all sequences containing those segments. The vowels in this data are not equivalent to steady state vowels and do not share the formant values presented in the previous acoustic analysis of English and Spanish monophthongs mentioned above.

5.6 Conclusion

The results show that the three expected alternations of monophthong, diphthong, and hiatus are present in this data. Monophthongs make up a small part of the data, with most of the data split between diphthong and hiatus pronunciations. Duration proved to be a successful differentiator between diphthong and hiatus pronunciations, and trends were identified in the data. The results differed by task type. In the Question and Answer data sets, diphthongization increased with level of acquisition. The stimuli contained about 80% of words with diphthongs as their target pronunciation, according to
the gliding rule in Spanish. As the level of acquisition increased, the level of
diphthongization moved towards 80%, although this percentage was not reached by any
of the participant groups. However, the results from the Nonsense Words task showed
that each group had similar levels of diphthongization, with slightly more hiatus being
pronounced than diphthongs. In sum, the real word data confirms that L2 learners have
more success with diphthongization as their level of acquisition increases, which
confirms the results in MacLeod (2012). The Nonsense Words data does not show a
similar relationship. The frequency data does not show consistent results to support the
idea that vowel quality in vowel sequences changes with level of acquisition. However,
there are a few sequences that hint at such a relationship. The next chapter provides a
theoretical analysis of the data presented in this chapter. Optimality Theory’s ability to
account for the variability in L2 output, as demonstrated in this data set, is to be
determined in the next chapter.
Chapter 6: Optimal Diphthongs

6.1 Introduction

This chapter analyses the output presented in Chapter 5 in an Optimality Theoretic framework. To review from Chapter 3, OT (Prince & Smolensky 1993/2004) is a generative theory, used mostly for phonology, that is constraint-based rather than rule-based. OT’s main departure from rule-based accounts is that it considers a set of universal violable constraints rather than language specific rules. Rather than ordering rules, OT ranks universal constraints in a language specific hierarchy. In OT learnability models, such as the Gradual Learning Algorithm (GLA) (Boersma & Hayes 1999), constraints are ranked and re-ranked until the learner acquires the form.

This chapter begins with an analysis of the three expected alternations for pronouncing vowel sequences in Spanish, monophthongs, diphthongs, and hiatus. The analysis then moves on to alternate or unexpected pronunciations, which occurred mostly in the output of the L2 groups. The tenets of the GLA are applied to the data to account for the variability in output at the three different levels of acquisition. The continuous nature of the GLA also benefits an analysis of the variation in the Native Speaker group’s output. This chapter answers the question of whether OT learnability models can account for all L2 output. As this chapter will demonstrate, OT encounters limitations in accounting for all of the data in this study. Frequency based models are discussed as a viable alternative.

6.2 Monophthongs

This section analyzes the monophthong pronunciations of sequences of two vowels in Spanish. The motivation for monophthongization is to avoid onsetless
syllables when two vowels are in contact. Monophthongs resulted from deletion, reduction, or coalescence. See section 5.3 for the distribution of monophthongs in the data set as a whole and for the distribution of monophthongization processes. Each process is analyzed below.

6.2.1 Deletion

Deletion processes are well known to occur across word boundaries in Spanish. For example, /e/ is often deleted in casual speech when unstressed and in contact with another vowel. In Chicano Spanish, low vowels are deleted across word boundaries when in contact with a word-initial vowel. Also, when a high vowel follows a mid vowel across word boundaries, the mid vowel is deleted (Baković 2006).

English also shows a dispreference for two vowels in sequence. One of the processes used to avoid vowel sequences is contraction to eliminate one of the vowels, as occurs with the word ‘is’ (Stockwell & Bowen 1965: 108). Contraction of ‘is’ also occurs with word-final consonants. Contraction of other words such as ‘not’ and ‘will’ are other examples of vowel deletion in English.

Deletion is motivated by the desire to eliminate onsetless syllables. There is also a desire to avoid syllable codas when possible. Both of these conditions must be met without a change in vowel quality. Below is the constraint hierarchy for low vowel deletion that occurs in Chicano Spanish.
Fig. 6.1 Low vowel deletion (Colina 2009: 63)

low/µ, ONSET >> *CODA >> MAX-V >> IDENT(high), hi/µ

/la ixa la hija ‘the daughter’

<table>
<thead>
<tr>
<th></th>
<th>low/µ</th>
<th>ONSET</th>
<th>*CODA</th>
<th>MAX-V</th>
<th>IDENT(high)</th>
<th>hi/µ</th>
</tr>
</thead>
<tbody>
<tr>
<td>li.xa</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>la.i.xa</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lai.xa</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This figure shows the constraint ranking that ensures an onset and does not create a coda. Because /a/ does not glide in Spanish, another option would be to raise the vowel to a mid or high vowel that could glide to create a falling diphthong, which Chicano Spanish does not allow across word boundaries. Below is the constraint ranking for maintaining vowel quality.

Fig. 6.2 Maintain vowel quality (Colina 2009: 63)

IDENT(low) >> MAX-V >> IDENT(high), hi/µ

<table>
<thead>
<tr>
<th></th>
<th>IDENT(low)</th>
<th>MAX-V</th>
<th>IDENT(high)</th>
<th>hi/µ</th>
</tr>
</thead>
<tbody>
<tr>
<td>li.xa</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lëi.xa</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lëi.xa</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

These two constraint rankings demonstrate why deletion is the preferred outcome, but they do not explain why the low vowel is deleted over the high vowel. In contrast to the Chicano data, the high vowel was deleted most frequently in the present study, regardless of its position in the sequence. Low and mid vowels are preferred nuclei over high vowels because they are more sonorous. The constraint MAX-V does not differentiate between the high and non-high vowels in the input, and Colina (2009) and Baković (2006) do not include a candidate with the low vowel deleted. A candidate with low vowel deletion would fair just as well under the constraint ranking above. MAX is
not sufficient for differentiating between the two possible outcomes of deletion (Casali 1997: 500). MAX-V must be modified to allow for the distinction between deleting high and non-high vowels. The constraints offered below will do the job.

**Fig. 6.3 MAX-V constraints for deletion**

- **MAX-V [-high]:** a \( V_{[\text{-high}]} \) in the input must have a corresponding \( V_{[\text{-high}]} \) in the output
- **MAX-V [+high]:** a \( V_{[+\text{high}]} \) in the input must have a corresponding \( V_{[+\text{high}]} \) in the output

**Fig. 6.4 Delete high vowel**

\[ \text{ONSET} >> \text{IDENT}(\text{high/low}) >> \text{MAX-V}_{[-\text{high}]} >> \text{hi}/\mu >> \text{MAX-V}_{[+\text{high}]} \]

<table>
<thead>
<tr>
<th>/tieko/ ‘tieco’ [te.ko] nonsense word</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{te.ko} )</td>
</tr>
<tr>
<td>( \text{ti.ko} )</td>
</tr>
<tr>
<td>( \text{tje.ko} )</td>
</tr>
<tr>
<td>( \text{ti.e.ko} )</td>
</tr>
<tr>
<td>( \text{ta.ko} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>/seis/ ‘seis’ six [ses]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{ses} )</td>
</tr>
<tr>
<td>( \text{sis} )</td>
</tr>
<tr>
<td>( \text{seis} )</td>
</tr>
<tr>
<td>( \text{se.is} )</td>
</tr>
<tr>
<td>( \text{sas} )</td>
</tr>
</tbody>
</table>

The above constraint ranking correctly predicts that deletion is the optimal strategy for some speakers for resolving hiatus when there is a high vowel and a non-high vowel in sequence, although not necessarily in that order. It also correctly predicts that the low vowel will be deleted. The speakers’ motivation in these pronunciations is to satisfy onset, but it is also to create the best nucleus. The most unmarked nucleus is a low vowel, so speakers are choosing the best nucleus from the available input.
However, some participants maintained the high vowel instead of the low vowel. The motivation for maintaining the high vowel is unclear. The high vowel was maintained infrequently and only in sequences of rising sonority, which means that the first vowel in the sequence was deleted. The first vowel was also deleted most frequently in sequences of two high vowels.

One option for explaining deletion in sequences of high and non-high vowels is the constraint ranking below, which is similar to the ranking in Fig. 6.5. In the constraint ranking below, MAX-V[+high] is ranked above MAX-V[-high] to ensure that the high vowel is maintained instead of the non-high vowel.

**Fig. 6.5 Delete non-high vowel**

\[
\text{ONSET} \gg \text{MAX-V[+high]} \gg \text{hi/μ} \gg \text{MAX-V[-high]}
\]

/tieko/ ‘tieco’ [ti.ko] nonsense word

<table>
<thead>
<tr>
<th></th>
<th>ONSET</th>
<th>MAX-V[+high]</th>
<th>hi/μ</th>
<th>MAX-V[-high]</th>
</tr>
</thead>
<tbody>
<tr>
<td>te.ko</td>
<td></td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>ti.ko</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>ti.e.ko</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

The constraint ranking in Fig. 6.5 does not explain the case of deletion of the vowel in sequences of two high vowels, in which the first vowel is most commonly elided. Casali (1997) points out that the decision to delete one vowel over another in order to resolve hiatus is not random; rather, universal constraints determine which element is deleted. He notes that deletion of the first element in a vocalic sequence in cross-linguistically most common, but that elision of the second segment also occurs (Casali 1997: 493). Accounts of deletion as a form of hiatus resolution (Casali 1997, Jenkins 1999, Baković 2006) deal with sequences that occur between words, but the data
in this study occurs within the word. Hence, the analyses done by Casali (1997) and Jenkins (1999) that focus on sequential position are of limited use here, as they give prominence to the status of the two words in the sequence as function words or lexical words. Baković (2006) shows the prominence of word-initial position in opting for first vowel deletion, something that also does not apply to this data set.

As Casali notes, “[OT] provides no direct mechanism for simply stipulating which vowel is elided” (1997: 499-500). The ranking of the syllabic constraints below disfavors a diphthong or a hiatus pronunciation. While deletion is favored, it is unclear which vowel should be deleted, and the motivation for choosing one over the other is also unclear. The available constraints account for deletion based on vowel quality, sonority, or syllabic structure, but deletion based on the linear property of the segments cannot be motivated by the available constraints. The $\mathcal{E}^{-}$ icon indicates that the constraint ranking predicts an undesirable output.

**Fig. 6.6 Delete first vowel**

\[ \text{ONSET} \gg \ast \text{CODA} \gg \ast \text{COMPLEX NUCLEUS} \gg \text{MAX-V}_{[\text{+high}]} \]

<table>
<thead>
<tr>
<th>/siudad/ ‘ciudad’ [su.dad] city</th>
<th>ONSET</th>
<th>*CODA</th>
<th>*COMPLEX NUCLEUS</th>
<th>MAX-V_{[\text{+high}]}</th>
</tr>
</thead>
<tbody>
<tr>
<td>su.dad</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>si.dad</td>
<td>$\mathcal{E}^{-}$</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>sju.dad</td>
<td>*!</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>sju.dad</td>
<td>*!</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>si.u.dad</td>
<td>*!</td>
<td></td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

A similar situation holds for sequences of two non-high vowels. In these sequences, the second vowel elided most frequently, regardless of vowel quality. The ranking of the syllabic and moraic constraints again disfavors a diphthong or a hiatus. The constraint ranking favors deletion, but it is unclear as to which vowel should delete.
There does not appear to be anything in the vowel quality to hint at which vowel should be deleted. *CODA and *COMPLEX NUCLEUS are included when necessary to rule out diphthongization, depending on which vowel is more likely to glide.

**Fig. 6.7 Delete second vowel**

\[
\text{ONSET} >> *\text{COMPLEX NUCLEUS} >> \text{IDENT(high/low)} >> \text{mid/μ} >> \text{hi/μ} >> \text{MAX-V}
\]

/kreatiba/ ‘creativa’ [kre.ti.va] creative

<table>
<thead>
<tr>
<th></th>
<th>ONSET</th>
<th>*COMPLEX NUCLEUS</th>
<th>IDENT(high/low)</th>
<th>mid/μ</th>
<th>hi/μ</th>
<th>MAX-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>kre.ti.va</td>
<td>kre.ti.va</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kra.ti.va</td>
<td>kra.ti.va</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kre.a.tiva</td>
<td>kre.a.tiva</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

/paoto/ ‘paoto’ [pa.to] nonsense word

<table>
<thead>
<tr>
<th></th>
<th>ONSET</th>
<th>*CODA</th>
<th>IDENT(high/low)</th>
<th>mid/μ</th>
<th>hi/μ</th>
<th>MAX-V</th>
</tr>
</thead>
<tbody>
<tr>
<td>pa.to</td>
<td>pa.to</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>po.to</td>
<td>po.to</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pa.q.to</td>
<td>pa.q.to</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pa.u.to</td>
<td>pa.u.to</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pa.o.to</td>
<td>pa.o.to</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 6.2.2 Coalescence

Coalescence was another method used by participants to resolve hiatus. In Spanish, identical vowels contract both across word boundaries and word-internally, reducing their duration to that of a single vowel (Hualde 2005: 90-92). Baković (2006) also claims coalescence for similar vowels, not just identical vowels. The constraint UNIFORMITY is violated through coalescence.

**Fig. 6.8 Constraint and constraint ranking for coalescence**

UNIFORMITY: No coalescence (McCarthy & Prince 1995)

\[
\text{ONSET} >> \text{UNIFORMITY} \ (\text{Baković 2006; Colina 2009})
\]
The following tableau shows how this constraint ranking predicts coalescence with necessary constraints added for the particular input. Only the necessary constraints are included.

**Fig. 6.9 Coalesce vowels**

\(/pue_{12}ve\) ‘pueve’ [po.ve] nonsense word

<table>
<thead>
<tr>
<th></th>
<th>ONSET</th>
<th>*COMPLEX NUCLEUS</th>
<th>MAX-V</th>
<th>IDENT(high)</th>
<th>UNIFORMITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>po_1,2.ve</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>pu_1.ve</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>pe_2.ve</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>pue.ve</td>
<td></td>
<td></td>
<td></td>
<td>* !</td>
<td></td>
</tr>
<tr>
<td>pu.e.ve</td>
<td></td>
<td></td>
<td></td>
<td>* !</td>
<td></td>
</tr>
</tbody>
</table>

UNIFORMITY is low-ranked to allow for coalescence to occur. For those speakers that did not select coalescence as the optimal strategy for hiatus resolution, UNIFORMITY would be more highly ranked to ouster coalescence as an option in favor of having diphthongization or deletion satisfy ONSET.

### 6.2.3 Reduction

Between English and Spanish, vowel reduction is thought of as an English phonological process; however, it does occur in dialects of Spanish. For example, in dialects in which /s/ is maintained, unstressed vowels are often shortened or deleted, such as in Mexican and Andean dialects (Lipski 1994). In Andalusian Spanish, the mid vowels reduce when /s/ has been deleted, so that /e/ \(\rightarrow\) /ɛ/ and /o/ \(\rightarrow\) /ɔ/ (Hualde 2005).

Although vowel reduction does occur in dialects of Spanish, it is not a widespread process in the language.

As stated in Chapters 2 and 3, vowel reduction occurs in relation to stress in English. Unstressed vowels are generally reduced to [ə]. As would be expected based on comparative studies (Stockwell & Bowen 1965, Whitley 2002), L2 participants in this
study showed reduced pronunciations that are not part of the Spanish vowel inventory. Many of these pronunciations were realized through coalescence. Other examples are reduced forms of the vowel left after deletion. They occur in stressed and unstressed syllables. Burzio (2007) explains why sometimes unstressed syllables in English contain a full vowel. His account is an improvement on both rule-based and other constraint-based accounts that focus solely on stress. He uses an OT analysis based on perceptual distinctions of changes in vowel quality. The two most useful constraints for this analysis are defined below. E stands for energy, and Q stands for vowel quality.

**Fig. 6.10  Vowel reduction constraints** (Burzio 2007: 158)

- **DELTAD(E)**: Maximize the energy difference between stressed and unstressed vowels
- **IDENT Q**: Vowel quality in the input is the same, in terms of distinctive features or acoustic values, as in the output

The constraint ranking below predicts vowel reduction in an unstressed syllable in English.

**Fig. 6.11  Vowel reduction constraint ranking in English** (Burzio 2007: 159)

\[ \Delta E >> \text{IDENT Q} \]

/bibliote\text{"2\text{"} teka/ ‘biblioteca’ [bɪbl\text{"}a\text{"}ˌ teka] library

<table>
<thead>
<tr>
<th></th>
<th>ONSET</th>
<th>ΔE</th>
<th>*COMPLEX NUCLEUS</th>
<th>MAX-V</th>
<th>IDENT Q</th>
<th>UNIFORMITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>_sprites</td>
<td>bɪbl\text{&quot;}a\text{&quot;}ˌ teka</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>_sprites</td>
<td>bɪblo\text{&quot;} teka</td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_sprites</td>
<td>bɪblo\text{&quot;} teka</td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_sprites</td>
<td>bɪblo\text{&quot;} teka</td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_sprites</td>
<td>bɪblo\text{&quot;} teka</td>
<td></td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In English, it is more important to have a perceptual difference between stressed and unstressed vowels than it is to maintain vowel quality (Burzio 2007: 160). Hence,
English speakers replace “the target quality with the quality that is articulatorily neutral”, which is schwa (Burzio 2007: 157). The motivation for reducing the unstressed vowel is to suppress “perceptually ineffective articulatory effort” (Burzio 2007: 156).

Burzio (2007) compares English and Italian, which does not have widespread vowel reduction. Spanish is similar in that stressed and unstressed vowels are pronounced the same, regardless of stress. Hence, in Spanish the constraint ranking would be reversed.

**Fig. 6.12 Constraint ranking for vowel quality maintenance in Spanish**

\[
\text{IDENT Q} >> \Delta E
\]

<table>
<thead>
<tr>
<th>ONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>biblio’teka</td>
</tr>
<tr>
<td>biblio2’teka</td>
</tr>
<tr>
<td>biblio’teka</td>
</tr>
<tr>
<td>bibli.o’teka</td>
</tr>
</tbody>
</table>

In Spanish, it is more important to maintain vowel quality than it is to express a difference between stressed and unstressed vowels.

In addition to a schwa pronunciation of unstressed vowels, participants produced other lax vowels that are not part of the Spanish vowel inventory. These pronunciations occurred in stressed and unstressed syllables. The constraint violated in these cases is *WEAK-LAX. Violation of this constraint results in “compromised perceptibility of [the] tense/lax distinction” (Burzio 2007: 162). Burzio is also able to account for intermediary reductions that fall between a full vowel and schwa through the continuous valuation of \( \Delta E \). This continuous valuation is similar to constraint ranking along the sonority scale (see Chapter 3). Languages differ in how much reduction they allow. \( \Delta E \) in a smaller...
script designates a smaller change in articulatory energy. The variability in L2 vowel production in this study can be reflected in this way.

**Fig. 6.13 Variability in L2 vowel quality**

*\text{WEAK-LAX}: No “perceptually weak laxness” (Burzio 2007: 162)

\[\Delta E \gg \text{WEAK-LAX}\]

\[\text{/su\_e\_ter/ [ˈsɛ\_ɛ\_tɛr] ‘súêter’ sweater}\]

<table>
<thead>
<tr>
<th></th>
<th>ONSET</th>
<th>$\Delta E$</th>
<th>*WEAK-LAX</th>
<th>$\Delta E$</th>
<th>MAX-V</th>
<th>*COMPLEX NUCLEUS</th>
<th>IDENT Q</th>
<th>UNIFORMITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>sɛ_ɛ_ter</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>se_ter</td>
<td>!</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>su_eter</td>
<td>!</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

This constraint ranking predicts the reduction that occurs in stressed syllables, in which an intermediary lax vowel is realized. The fact that the unstressed vowel is not reduced does not complicate the analysis.

While this account does not follow the syllabic perspective that the rest of this analysis undertakes, it is a straightforward and well-motivated analysis that allows for cross-linguistic comparison of stressed and unstressed vowel inventories. Accounts such as Hammond’s (1999), as was stated in Chapter 3, have more difficulty in explaining why unstressed vowels are not always reduced. Hammond (1999) has to propose language specific constraints, and his analysis of vowel reduction in general is encumbered by the controversy over ambisyllabicity, which he defends using constraints that are typologically unmotivated, such as MAX CODA. The analysis presented here is free of such complications.
6.3 Diphthong-Hiatus Contrast

Participants most commonly alternated their pronunciations between diphthong and hiatus. This alternation occurs in dialects of Spanish, as described in Chapters 2 and 3. The constraint ranking for diphthongization in Spanish appears below, repeated from Chapter 3 with words from this data set. The motivation for diphthongization is to avoid onsetless syllables.

Fig. 6.14 Diphthongization (rising diphthongs only) (Colina 2009: 23)

\[ \text{*ONSET/glide, ONSET} \gg \text{*COMPLEX NUCLEUS, *NUC/glide, MAX-ΙΟμ} \]

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{/familia/ famil}ia & \text{‘family’} & \text{*ONSET/glide} & \text{ONSET} & \text{*COMPLEX NUCLEUS} & \text{*NUC/glide} & \text{MAX-ΙΟμ} \\
\hline
\text{a. } & \text{familia} & & & * & * & * \\
\text{b. famili.a} & & *! & & & & \\
\hline
\end{array}
\]

Fig. 6.15 Diphthongization (falling diphthongs) (Colina 2009: 42)

\[ \text{*ONSET/glide, ONSET} \gg \text{*CODA/glide, *CODA, MAX-ΙΟμ} \]

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{/seis/ seis} & \text{‘six’} & \text{*ONSET/glide} & \text{ONSET} & \text{*CODA/glide} & \text{*CODA} & \text{MAX-ΙΟμ} \\
\hline
\text{a. } & \text{seis} & & & * & * & * \\
\text{b. se.is} & & & *! & & & \\
\hline
\end{array}
\]

The constraint ranking for hiatus pronunciations would be a reversal of the rankings above since hiatus is a clear violation of ONSET. The hiatus pronunciation may also be motivated by the desire to avoid illicit consonant + glide sequences in English. CiV sequences only occur when the post-glide vowel is /u/. To account for English’s influence on this pronunciation, it is necessary to add constraints from Baertsch’s (2008) account that utilizes constraints based on the peak and margin hierarchies to explain permissible sequences. These constraints would all outrank ONSET, but would be parallel to the nucleus and mora constraints.
Fig. 6.16 Hiatus (rising sequence)

*COMPLEX NUCLEUS, *NUC/glide, MAX-IOμ >> *ONSET/glide, ONSET

<table>
<thead>
<tr>
<th>/familia/ familia ‘family’</th>
<th>*COMPLEX NUCLEUS</th>
<th>*NUC/glide</th>
<th>MAX-IOμ</th>
<th>*ONSET/glide</th>
<th>ONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. familja</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b. famil.i.a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 6.17 Hiatus (rising sequence): No illicit C-liV sequences

*M/L_i2 >> AGREE[hi] >> *P/i >> ONSET

<table>
<thead>
<tr>
<th>/familia/ familia ‘family’</th>
<th>*M/L_i2</th>
<th>AGREE[hi]</th>
<th>*P/i</th>
<th>ONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>famil1i.a</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>famil1i2a</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>famil1i.a</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 6.18 Hiatus (falling sequence)

*CODA/glide, *CODA, MAX-IOμ >> *ONSET/glide, ONSET

<table>
<thead>
<tr>
<th>/seis/ seis ‘six’</th>
<th>*CODA/glide</th>
<th>*CODA</th>
<th>MAX-IOμ</th>
<th>*ONSET/glide</th>
<th>ONSET</th>
</tr>
</thead>
<tbody>
<tr>
<td>seis</td>
<td>*!</td>
<td>**</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>sein</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The relative ranking of the coda and onset constraints along with the ranking of the mora faithfulness constraint determine the diphthong or hiatus pronunciation. These rankings predict the largest percentage of this study’s data.

6.4 Other Pronunciations

There were several other pronunciations aside from the monophthong-diphthong-hiatus alternation. Participants in the L2 groups produced most of these tokens. See section 5.4.2 for examples of these pronunciations and their distribution. Each pronunciation is analyzed below.
6.4.1 Metathesis

In a few examples, participants metathesized the two vowels in the vowel sequence. This process was carried out in the sequences /ai/, /ei/, and /iu/. The motivation for this pronunciation of the two rising diphthongs was to satisfy ONSET, but also to have the least information possible in the coda by moving the segment most likely to glide to a prevocalic position. The constraint LINEARITY is a faithfulness constraint that blocks metathesis (McCarthy & Prince 1995). In other words, the order of the segments must be maintained. Below is the constraint ranking that predicts metathesis. Only the necessary constraints are shown.

**Fig. 6.19 Constraint ranking for metathesis**

LINEARITY-IO: the precedence of segments in the input must be the same in the output (McCarthy & Prince 1995)

ONSET >> *CODA, *COMPLEX CODA >> LINEARITY-IO

In the above example, ONSET was satisfied while retaining as little information in the coda as possible. However, metathesized sequences were also pronounced in hiatus. The motivation for the hiatus pronunciation may be to bar against an impermissible onset cluster in English. Baertsch (2008) explains the distinct parsings for iV sequences based on Prince & Smolensky’s (1993/2004) peak and margin hierarchies, which determine acceptable segments for each part of the syllable based on the sonority scale (see Chapter 3).
In the above example with the vowels in /seis/ metathesized, the English ranking would keep the vowels in hiatus so that /seis/ would be [si.es]. In fact, this constraint ranking has a part in explaining hiatus pronunciations of all iV sequences in this data set. The constraint ranking is below, with LINEARITY-IO added to account for the violation caused by metathesis.

**Fig. 6.20 Constraint ranking for CiV(other than /u/) sequences**  
(Baertsch 2008: 236)

<table>
<thead>
<tr>
<th>/seis/ [si.es] six</th>
<th>*Obs₁i₂</th>
<th>AGREE[hi]</th>
<th>*P/i</th>
<th>ONSET</th>
<th>LIN-IO</th>
<th>*P/-hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>s₁i.es</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>s₁j2es</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>s₁jjes</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

The above constraint hierarchy also explains the switch for /iu/ sequences. Glided /u/ is permissible after many consonants in English, and the vowels that can follow it are not limited, as is the case with /i/. For example, participants metathesized the vowels in *ciudad* ‘city’ and *tuisa* (nonsense word). The pronunciation may have been motivated by a desire to block an impermissible onset. Both [sw] and [tw] are permissible onsets in English, but [sy] and [ty] are not (Davis & Hammond 1995). Only the necessary constraints are included.

**Fig. 6.21 [iu] → [ui] metathesis**  
(Baertsch 2008: 231)

<table>
<thead>
<tr>
<th>/siudad/ ciudad ‘city’ [suiidad]</th>
<th>*P/u</th>
<th>*Obs₁u₂</th>
<th>*P/i</th>
<th>ONSET</th>
<th>*M₂/i</th>
<th>LIN-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>s₁uiedad</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>s₁uiedad</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>s₁uiedad</td>
<td>*!</td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>s₁u.idad</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
6.4.2 High Vowel Displacement

High vowel displacement occurred when participants moved the high vowel to another syllable in the word. This pronunciation is similar to metathesis in that there is a change in the order of the segments. Just like metathesis, this change would involve a violation of LINEARITY. This pronunciation occurred with /i/. For example, agaico → agacio, goiro → gorío. This change produced both diphthongs and hiatus. The motivation for this pronunciation over other strategies for hiatus resolutions in unclear. This change could also result from analogy based on common Spanish words that end in –io or –ío.

Fig. 6.22 /i/ displacement

*CODA >> ONSET >> LINEARITY

<table>
<thead>
<tr>
<th>/agaiko/ [agasio] (nonsense word)</th>
<th>*CODA</th>
<th>ONSET</th>
<th>LINEARITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>agasio</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>agaiko</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>agasi.o</td>
<td></td>
<td>*!</td>
<td></td>
</tr>
</tbody>
</table>

To rule out the form with high vowel displacement and diphthongization, it is necessary to include *COMPLEX NUCLEUS and rank it above ONSET.

Fig. 6.23 /i/ displacement

<table>
<thead>
<tr>
<th>/goiro/ [go′ rio] (nonsense word)</th>
<th>*CODA</th>
<th>*COMPLEX NUCLEUS</th>
<th>ONSET</th>
<th>LINEARITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>go′ rio</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>goi.ro</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>go.rio</td>
<td></td>
<td>*!</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
6.4.3 High Vowel Insertion

/ɪ/ was inserted in the last syllable of some of the nonsense words. For example, agaico → agaicio, tuisa → tuisia. Insertion is a violation of DEP. These could have resulted from analogy based on other Spanish words that commonly end in –io or –ia.

6.4.4 Change in Vowel Quality

Change in vowel quality sometimes happened with cognates and reflected English pronunciation, such as biología → b[aj.o]logía. However, most changes in vowel quality had little to do with the English pronunciation. These changes were not reductions; rather, they were changes to other vowels in the Spanish vowel inventory. The changes occurred in stressed syllables. All of these changes involve at least one violation of IDENTQ, which is low ranked. Some of the changes require a change in syllabic structure, such as the one mentioned above, which creates a hiatus and adds a coda segment. Other changes make no changes to syllabic structure, such as t[e.a]tro → t[i.e]tro.

6.4.5 Consonant Insertion

Consonants were inserted between vowels in a few examples, such as daeta → dalita. The most frequent consonant inserted between vowels was a glottal stop. Consonant insertion is a clear attempt to create an onset and is a clear violation of DEP. Included in this process is the insertion of [j] before /u/ in the word universidad ‘university’. This pronunciation shows influence from English, but it can also be considered an effort to satisfy ONSET.
6.24 Consonant insertion

ONSET >> *CODA >> DEP

/daeta/ [da.li.ta] (nonsense word)

<table>
<thead>
<tr>
<th></th>
<th>ONSET</th>
<th>*CODA</th>
<th>DEP</th>
</tr>
</thead>
<tbody>
<tr>
<td>da.e.ta</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>da.li.ta</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>dal.i.ta</td>
<td>*!</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

6.4.6 Summary

The common motivation for all of the pronunciations mentioned above is to satisfy ONSET through some manner of hiatus resolution. One way to satisfy ONSET was to create a monophthong through deletion or coalescence. In both of these processes, it was more desirable to have an unmarked output and to allow violations of the faithfulness constraints of IDENT-V, MAX-V, and UNIFORMITY.

Diphthongization was the most commonly used form of hiatus resolution in the data set. In the case of rising sequences, ONSET dominated constraints that barred complex nuclei and the faithfulness constraint against deleting a mora from a high vocoid. In the case of falling diphthongs, ONSET dominated coda constraints. When hiatus was the preferred pronunciation, the constraints for nucleus and coda, along with the moraic faithfulness constraint, dominate ONSET. Speakers must decide if it is more desirable to have an unmarked nucleus or coda, as is the case with hiatus, or an unmarked onset, as is the case with diphthongization. Faithfulness to moraic structure also plays a role in this decision. Maintaining hiatus is also motivated by the desire to avoid illicit C\i\i\V sequences in English.

There were other strategies that participants used less frequently than monophthongization and diphthongization. Each of these pronunciations also places
primacy on satisfying ONSET. Both metathesis and the displacement of /i/ to another syllable satisfy ONSET at the expense of LINEARITY. Both consonant insertion and word-initial /i/ insertion satisfy ONSET at the expense of DEP.

Influence from English constraint rankings was most directly observed in vowel reduction. The analysis for vowel reduction broke from the syllabic perspective adopted in the rest of this analysis, but it allowed for useful cross-linguistic comparison. In English, it is more important to create a distinction in vowel quality between stressed and unstressed syllables than it is to maintain vowel quality. In Spanish, the opposite holds true. Faithfulness, in the form of IDENT-Q, is ranked lower in English than it is in Spanish when it comes to vowel pronunciation.

Stockwell & Bowen argue that the one strategy that English speakers revert to most frequently is insertion of a glottal stop to resolve hiatus (1965: 110). However, the data here shows that English speakers use a variety of options to break up vocalic sequences. There is also evidence of analogy to common Spanish words to which the L2 learners may have already been exposed and to English pronunciations of sequences in cognates or sequences found in common English words. In sum, diphthongization was the most commonly used strategy for resolving hiatus, but the rate at which diphthongization and other processes were used between the different participant groups has not been explained. The next section takes up this topic.

6.5 Gradual Learning Algorithm

The GLA (Boersma & Hayes 1999) is a learnability model within the framework of OT. This algorithm has been able to account for both L1 and L2 acquisition (see Chapter 3) through the use of continuous constraint ranking, stochastic evaluation, and
through allowing for optionality in the output and different levels of well-formedness.

This section will demonstrate how the tenets of the GLA can account for the data presented in Chapter 5 using the constraint rankings presented above.

6.5.1 Diphthong-Hiatus Contrast

As the results in Chapter 5 show, diphthong production in the real word data from the Question & Answer task increased with level of acquisition towards the Native Speaker percentages, while hiatus production decreased. The results are repeated here from Chapter 5.

Fig. 6.25 Results by task

<table>
<thead>
<tr>
<th>Sound</th>
<th>Result</th>
<th>Question</th>
<th></th>
<th>Answer</th>
<th></th>
<th>Nonsense</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>I</td>
<td>A</td>
<td>N</td>
<td>B</td>
</tr>
<tr>
<td>MONOPHTHONG</td>
<td></td>
<td>3%</td>
<td>2%</td>
<td>1%</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>DIPHTHONG</td>
<td></td>
<td>25%</td>
<td>50%</td>
<td>55%</td>
<td>66%</td>
<td>31%</td>
</tr>
<tr>
<td>HIATUS</td>
<td></td>
<td>69%</td>
<td>49%</td>
<td>44%</td>
<td>32%</td>
<td>64%</td>
</tr>
</tbody>
</table>

In the GLA, constraints have ranges to allow for a noise component at evaluation time. At the evaluation point, the learner can make a selection anywhere within the range. Because the constraints’ ranges can overlap, constraint ranking may vary depending on the selection point. In this way, learner outputs will vary until converging on Native Speaker norms.

The following figures show how the constraints interact along a continuum of strictness and laxness for diphthongization in Spanish. Diphthongization is the result of the interaction between ONSET and CODA or NUCLEUS (depending of the rising or falling nature of the diphthong) constraints. MAX-ḫ constraints would align with the NUC/CODA constraints, but for ease of presentation, they will not be included below. At all levels of
acquisition, the constraint groups will overlap, but the amount of overlap and the order of the constraints will change with level of acquisition.

**Fig. 6.26 Constraint ranking by level of acquisition: Beginner group**

If the selection points occur at non-overlapping points in the constraint range, then it is clear that NUC/CODA dominates ONSET, resulting in a hiatus. However, it is possible to select points at which the opposite ranking is true. For example, C₃ is a high point in the ONSET range that outranks a low point in the NUC/CODA range. If C₃ and C₄ are the selection points at the moment of evaluation, then a diphthong will be the optimal output. This model reflects that hiatus is the most frequent output in this participant group but that it is possible for diphthongs to surface.

In the Intermediate group, there were more diphthongs than hiatus. This relationship is reflected in the figure below in which ONSET is more highly ranked than NUC/CODA. However, diphthongs occur only slightly more frequently than hiatus; hence, the overlap between the two constraints on the strict/lax continuum is greater than that of the overlap between the constraints for the Beginner group in which the percentages were more disparate.
Because the overlapping area of the constraint ranges is great, there is less distinction in the probability that one constraint will outrank the other. The non-overlapping portions represent the small percentage difference between diphthongs and hiatus in which ONSET dominates NUC/CODA to account for slightly more diphthongs than hiatus.

The Advanced group had more diphthongs than hiatus, so that the overlap between the constraint ranges will be smaller than that of the Intermediate group.

The non-overlapping areas are greater than the overlapping area, reflecting the fact that there are more instances in which ONSET unequivocally dominates NUC/CODA, meaning that more diphthongs occur than hiatus.

In the Native Speaker group, there is little overlap between the constraint ranges, as diphthong and hiatus pronunciations were more or less at expected levels based on the constraint hierarchies for diphthongization in Spanish.
For Native Speakers, the noise component at evaluation is quite low, as we find that
diphthongization is the optimum strategy for hiatus resolution.

Hiatus maintenance is motivated in the L2 groups by a desire to avoid CjV
sequences that are illicit in English. The constraints concerned with this avoidance are
highly ranked and would overlap to varying degrees in different levels of acquisition, the
overlap being greater in the Beginner group and decreasing as level of acquisition
increases. In the Native Speaker group, these constraints would be dominated. The
figure below demonstrates this ranking for the Beginner and Native Speaker groups. For
ease of presentation, the constraints blocking illicit CjV sequences will be represented as
*CjV.
6.5.2 Monophthongs and Other Pronunciations

The GLA is useful for accounting for other hiatus resolution strategies. After diphthongs and hiatus, monophthongs accounted for the next largest group of pronunciations. Each participant group pronounced monophthongs, but they were most common in the Beginner group. Monophthongs occurred as a result of deletion or coalescence, which constitute violations of MAX-V and UNIFORMITY, respectively. ONSET is highly ranked in monophthongs, and MAX-V and UNIFORMITY are low-ranked. However, monophthongs account for a small amount of the total data, so that MAX-V and UNIFORMITY must be more highly ranked than the NUC/CODA constraints for most of the output. Because the elided and the more faithful outputs are in competition, there must be a small amount of overlap to predict outputs with deletion as the optimal candidate over diphthongized or completely faithful forms. See this ranking in the figure below. For the other participant groups, the ranking would be the same.
except that the overlapping area would be smaller as the occurrences of monophthongization decrease.

**Fig. 6.31 Deletion or coalescence: Beginner group**

![Diagram](max-v/unif-strict-lax)

The other hiatus resolution strategies that OT accounts for in the data set are metathesis and consonant insertion. These strategies satisfy ONSET and violate LINEARITY, in the case of metathesis, and DEP, in the case of consonant insertion. These strategies were found in the L2 groups and were seldom used by the Native Speaker group. As these strategies are in competition with monophthongization, diphthongization and hiatus NUC/CODA constraints, a small amount of overlap is expected to account for the low percentage of these pronunciations in the data set. In the Native Speaker data, there would be little to no overlap because LINEARITY and DEP would be highly ranked to block metathesis and insertion in favor of diphthongization.

OT and the GLA account for much of the data, both in the L2 groups and the Native Speaker group. The GLA is able to account for the variation in output forms and the variation between levels of acquisition and the Native Speaker group. However, OT cannot fully account for some of the pronunciations. Deletion in sequences of two high vowels and in sequences of two non-high vowels is not explained, except that it is a MAX violation. OT cannot explain why one vowel was selected for deletion over the other in these sequences. OT does not explain pronunciations that result from analogy based on
English words or common Spanish words. The next section discusses frequency-based theories that may be able to account for some of this unexplained data.

6.6 Frequency

As stated above, OT is able to account for much of the L2 output reported in this study, but some of the data is left unexplained. One factor that has not been discussed in the analysis thus far is the difference in the data resulting from task type. Chapter 5 and Fig. 6.25 present a difference in the data resulting from the real word tasks and the data resulting from the Nonsense Word task. The difference is that while the real word data showed an increase in diphthongs over hiatus with an increase in level of acquisition towards the Native Speaker values, the Nonsense Word data showed little difference between diphthong and hiatus pronunciation based on level of acquisition. In fact, all groups, including the Native Speaker group were quite similar. OT and the GLA do not respond to this difference based on task type. Frequency based theories may be better able to account for this difference. The GLA does consider frequency in that more frequent input leads to the learner acquiring the correct form; however, frequency does not pertain to a particular lexical item in the GLA, but rather to a sequence of segments or morphemes. Because the Nonsense Words task presents a novel lexical set, it is necessary to consider theories of lexical frequency in explaining the difference in results in the present data set.

Bybee (2001) proposes a usage based model of phonology. Central to this model is the idea that humans categorize words as they do other types of information, so that phonological knowledge is not innate. Rather, phonological knowledge is determined by the input, and semantic information matters to the categorization of phonological
knowledge. Words are organized by context of use and phonetic similarity into schemas, which would not exist without the existence of the lexical units.

Bybee’s (2001) proposal embraces some aspects of OT, such as the violability of constraints, but it distances itself from many other tenets of OT. In OT, there is a separation between the lexicon and the grammar. While words and morphemes are essential units in generative phonological analyses, they are stripped of their semantic information. However, children learn sound sequences as they exist in words, not in isolation.

As in the GLA, frequency is a key element in Bybee’s (2001) model. The frequency of a word’s use influences the phonological structure of the word. Hence, there is a relationship between experience and representation. Through repetition, words are more easily accessed and their production becomes automatic. Nonce word acceptability is based on probability, but it is not entirely predictable. Schemas are gradient, meaning that there is a probability for one nonce word to apply to one category or another. This gradience is similar to the constraint ranges and continuous ranking in the GLA, but OT cannot explain why words with different frequencies behave differently.

There is a certain amount of unpredictability with nonce words. Even frequent patterns are not always passed on to nonce words (Bybee 2001). Bybee & Pardo (1981) study the alternation between diphthongs and mid vowels in stem-changing verbs in the Spanish of adult native speakers. They use nonce words that are patterned on existing verbs to gain an understanding of morphophonemic alternations. They find that the phonological rules applied to real verbs based on stress and vowel quality do not apply to
the nonce words in the same way. Hence, they conclude that the phonological rule is tied to the lexical item and that the rule does not exist independently in the grammar. Carlson (2006) used novel words to investigate the diphthong/mid vowel alternation in adult L2 learners of Spanish. He found that the distribution of diphthongs in the L2 learners’ speech was not the same as in the Spanish lexicon. Carlson (2010) found that L2 learners are sensitive to probabilistic variation in lexical naming tasks with nonce words, but that their sensitivity differs from that of native speakers.

Another central element of Bybee’s (2001) proposal is the concept of emergence, which is tied to the definition of universals in this model. “The units of phonology – syllables and segments – emerge from the interaction of the properties of gestural coordination and the frequency with which these gestures are combined sequentially” (Bybee 2001: 34). She defines universals as such: “The true universals of language are the dynamic mechanisms that cause language to change in certain systematic ways as it is used and as it is transmitted to new generations” (Bybee 2001: 189). In short, grammar emerges from the system in which it is used. Because Bybee advocates for a usage based linguistic model, she promotes the idea that “theories of competence should be moving closer and closer to plausible theories of performance, but [OT] seems to move in the opposite direction, as the style of derivation is psycholinguistically implausible in the extreme” (2001: 194).

6.7 Conclusion

This chapter presented an OT analysis to account for the data presented in Chapter 5. While OT is able to account for much of the data, there are some forms left unexplained. The GLA was also presented, and it was shown that its basic tenets can
account for the variability in the output. However, the difference between the real word tasks and the data from the Nonsense Word task could not be accounted for in an OT model. Usage based lexical frequency models, such as proposed by Bybee (2001), may be able to better respond to the different data presented in the nonsense words. It would be necessary to run further experiments and better control for frequency of two vowel sequences in the Spanish lexicon to definitively propose a usage based model of the acquisition of Spanish diphthongs.
Chapter 7: Conclusion

7.1 Summary

This study analyzed the production of sequences of two vowels in Spanish by adult native speakers of English in an OT framework. The overall goal was to gain insight into the L2 pronunciation of vowel sequences in Spanish, given the small amount of attention paid to this aspect of both vowel pronunciation and syllabification in the field of L2 phonology. An additional goal was to ascertain the extent to which OT learnability models can account for L2 production.

Previous research based on derivational phonology was presented to describe differences between the English and Spanish vowel systems and syllable structures. The SLA literature based on these derivational accounts demonstrates that the differences between the two languages do provide a source of difficulty for English speakers learning Spanish. The literature also shows that L1 transfer does not account for all L2 errors, and that universal markedness principles are evident in L2 phonological acquisition.

OT accounts of Spanish and English syllabification were also reviewed. Satisfying ONSET was found to be a motivating factor in both Spanish and English syllabification. Accounts of both languages also demonstrate how sonority can be formally incorporated into an OT analysis, something that is not possible in derivational accounts. The SLA literature based on OT learnability models was reviewed, which showed that OT is able to account for L2 speech production. However, some questions did surface as to the possible demotion of faithfulness constraints and as to how frequency is measured in the GLA.
64 participants provided the data for this study by completing two production tasks, which were a Nonsense Word task and a Question & Answer task. The duration of vowel sequences and the F2 values at various points along the vowel sequence’s duration were measured in order to categorize the sequences as monophthongs, diphthongs, or hiatus.

The data was then analyzed in an OT framework. The tenets of the GLA were then used to explain how the constraints interact to produce variation in the production of vowel sequences in L2 Spanish. A model based on lexical frequency was also discussed as an alternative to an OT account of L2 speech production.

7.2 Conclusions

This study was able to answer the research questions posed in Chapter 1. The questions are repeated here:

1) How do adult native speakers of English pronounce sequences of two vowels in L2 Spanish at different levels of acquisition?

2) Can OT learnability models, specifically the GLA, account for the pronunciation of L2 diphthongs? If so, what constraints do learners use and how do these constraints interact? If not, what other model(s) might offer an improved analysis of L2 diphthongs?

To answer the first question, it is necessary to consider the results from the two tasks separately. The data resulting from the Question & Answer task reveals that the rate of diphthongization increases with level of acquisition, approximating the level of diphthongization for the Native Speaker group. The stimuli in this task consisted of real words embedded in questions to which the participants responded.
The data resulting from the Nonsense Words task revealed little difference between the levels of acquisition. Furthermore, the L2 groups and the Native Speaker group had nearly equal levels of diphthong and hiatus pronunciations. The Nonsense Word task revealed no effect for level of acquisition.

Monophthongs were the third most common pronunciation, produced by deletion or coalescence. Some other strategies that were used mostly by L2 learners were metathesis, vowel reduction, /i/ displacement, /i/ insertion, and consonant insertion.

An OT analysis was conducted to answer the second question. The OT analysis was able to account for many of these pronunciations through constraints that have been established in the literature. The motivation for almost all of these pronunciations was to satisfy ONSET and resolve hiatus. However, OT was not able to account for the deletion of one vowel over another in sequences of vowels with the same [high] value. In these deletions, the choice was made based on the position of the vowel in the sequence, so there was no apparent phonological motivation that could be expressed through constraint interaction. Neither /i/ displacement nor /i/ insertion was used to create more optimal syllables. While it is easy to point to which constraints /i/ displacement and /i/ insertion violate (LINEARITY and DEP, respectively), it seems that they are motivated by analogy to common English or Spanish words. In addition, the vowel quality was changed, either to accommodate an English pronunciation or for some unexplained reason. In either case, a better syllable does not result. For the pronunciations that were accounted for in the OT analysis, it was shown that the tenets of the GLA do explain how this learnability model could account for the variation in L2 output.
This study concludes that level of acquisition is a factor in diphthong production in L2 Spanish. When confronted with a sequence of two vowels, Beginner learners favor hiatus pronunciation, while Advanced learners favor diphthongs. Other strategies for hiatus resolution were also used. This study also concludes that while OT can account for much of L2 speech production reported here, it is unable to explain all of the strategies used by L2 learners. The frequency factor in the GLA points in the direction of exploring usage-based lexical frequency models of language change (Bybee 2001) and how they can be used in SLA research.

7.3 Recommendations for Future research

Due to the scarce amount of research on diphthongization in L2 Spanish, both production and perception studies are needed to confirm the results in this study and others (Zárate-Sández 2011, MacLeod 2012). There is also a need for more acoustic studies of diphthongs in Spanish and English, as pronounced by native speakers of those languages. In the data available to date, diphthongs are rather elusive and it has been difficult for researchers to establish phonetic parameters in the same way that such parameters have been established for vowels and other areas of phonotactics.

If frequency-based methods are to be used, it is necessary to control better for frequency. This study obtained the frequency of vowel sequences using a frequency dictionary of Spanish (Juillard & Chang-Rodriguez 1964). The data in this dictionary is based on written material that is now dated by more than 50 years. The *Corpus del español* (Davies 2002) would provide more up to date data and is more widely used. In addition, it is suggested that perceived frequency of L2 learners be taken into account when constructing stimuli, following Broselow & Xu (2004).
This study did not control for the consonants adjacent to the vowel sequences. However, because there are illicit C\text{"}{\textupsilon}\ sequences in English, glide pronunciation was conditioned by the preceding consonant in the L2 data. Hence, it would be wise to control for types of consonants that precede high vowels in the stimuli, considering both frequency and articulation. Liquids and fricatives create resonances that can interfere with establishing boundaries for vowel segments. Depending on the aims of the particular study, it may be necessary to avoid these segments.

While this dissertation does not offer a detailed review of a usage-based model, such as that proposed by Bybee (2001), it does indicate that it is a promising avenue for SLA research that might be able to answer some questions remaining from the OT analysis. The results in this study concerning nonsense words are not surprising in light of lexical frequency based research (Bybee & Pardo 1981; Carlson 2006, 2010). Future studies containing repetition and stimuli that are more carefully controlled for frequency are needed to confirm the results in this study.


Berent, Iris, Donca Steriade, Tracy Lennertz, & Vered Vaknin. 2007. What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104: 591-630.


Levelt, Clara C. & Ruben van de Vijver. 2004. Syllable types in cross-linguistic and
developmental grammars. In René Kager, Joe Pater & Wim Zonneveld (eds.),
_Constraints in phonological acquisition_, 204-218. Cambridge, UK: Cambridge
University Press.


MacLeod, Bethany. 2012. Investigating L2 acquisition of Spanish vocalic sequences.


Martínez Celdrán, Eugenio. 1995. En torno a las vocales del español: Análisis y
reconocimiento. In _Estudios de fonética experimental_ 7: 197-218. Barcelona:
Promociones y Publicaciones Universitarias.

Martínez-Gil, Fernando & Sonia Colina (eds.), _Optimality-Theoretic studies in Spanish

University Press.


McCarthy, John J. & Alan S. Prince. 1994. The emergence of the unmarked: Optimality
in prosodic morphology. _Proceeding from the North East Linguistics Society_ 24:
333-379.

Beckman, Laura Walsh, & Suzanne Urbanczyk (eds.), _University of
Amherst, MA: Graduate Linguistic Student Association. (Available on Rutgers
Optimality Archive – ROA 60)


Appendix A

Consent to Participate in Research

Participant Consent Form:
Optimal Diphthongs: An Optimality Theoretic Analysis of the Acquisition of Spanish Diphthongs

PURPOSE:
My name is Alice Krause, and I am a doctoral student in the Department of Languages, Literatures & Cultures. I am doing research for my doctoral dissertation on the acquisition of Spanish diphthongs by native English speakers.

PROCEDURE:
You will be asked to pronounce a series of nonsense words, read and respond to a series of questions in Spanish, and speak briefly in Spanish about your daily routine. You will be audio recorded on a digital voice recorder in HU 286, which is a small conference room that I will reserve for the purposes of this study. You will also need to fill out a brief questionnaire about your background in Spanish. It should take no more than one half hour session to fill out the questionnaire and record your performance of the above mentioned tasks.

RISKS:
You may experience embarrassment upon speaking Spanish and/or being recorded. There is also a chance that recording could take longer than expected if technical difficulties arise.

BENEFITS:
You will not receive any direct benefits from participating in this research. You will not receive any extra credit in your Spanish class or any other class offered at the University at Albany in return for your participation in this study. However, your participation may be useful in informing future research in the field of second language acquisition.

CONFIDENTIALITY:
All information is confidential and no one but me will have access to it. I will be the only person in possession of your recording, and your recording will be erased from the digital recorder and my computer after the research has concluded. All electronic data will be stored on the digital recorder or as a file on my computer. This consent form and the questionnaire will be securely stored in a file cabinet in my office (HU 282). Although I am the only person who will have access to your recording, my supervising professor may be asked to listen to your recording in order to assist me in linguistic analysis. You are welcome to listen to your own recording. Your name will only be required when you sign this consent form if and when you decide to participate in the study. I will not ask for your name at any point after I start recording. Your background information will be connected to your recorded audio file via a randomly assigned number, which will be assigned just before recording begins. Your name will not appear on any recordings or electronic files. If at any time you wish to discontinue your participation in the study all data will be destroyed.

PARTICIPATION AND WITHDRAWAL:
Your participation in this study is entirely voluntary. Even after you agree to participate in the research or sign the informed consent document, you may decide to leave the study at any time without penalty, at which time, your data will be destroyed. Your participation in this research can be terminated by me without your consent in the following circumstances: you are a native
speaker of a language other than English or Spanish; you provided false/incorrect information on
the background questionnaire and you are unavailable to provide corrections; technical
difficulties arise that render your recording difficult or impossible to use and you are unavailable
to repeat the tests. It is possible that other unforeseen difficulties may arise that could result in
termination of your participation in this research without your consent.

IDENTIFICATION OF RESEARCHER:
If you have any questions or concerns please contact me, Alice Krause, at:
Phone: 518-442-4112
e-mail: ak4168@albany.edu

Or contact my supervising professor, Maurice Westmoreland, at:
Phone: 518-442-4246
e-mail: mw908@albany.edu

RIGHTS OF THE PARTICIPANTS:
You may withdraw your consent at any time and discontinue participation in this research. One
copy of this document will be kept with the research records of this study. Also, you will be given
a copy to keep.
If you have any questions concerning your rights as a research participant that have not been
answered by the investigator or if you wish to report any concerns about the study, you may
contact the University at Albany Office of Regulatory Research Compliance at 518-442-9050 or
orrc@uamail.albany.edu.

Please answer the following question before signing for your participation:

Do you agree to be audio recorded? Circle one: YES NO

Signature of Research Participant

I have read, or been informed of, the information about this study. I hereby consent to participate
in the study.

Name ________________________________

Signature ______________________________

E-mail ________________________________

Date ____________________
Appendix B

Documento de consentimiento

Optimal Diphthongs: An Optimality Theoretic Analysis of the Acquisition of Spanish Diphthongs

PROPÓSITO:
Me llamo Alice Krause y soy estudiante doctoral en el departamento de Languages, Literatures & Cultures. Hago investigación para mi tesis doctoral sobre la adquisición de diptongos en español de hablantes nativos de inglés.

MÉTODO:
Como participante, se le va a pedir que usted pronuncie una serie de palabras que no existen en español pero que se crean a base de español. Además, se le va a pedir que lea y responda a una serie de preguntas en español y que hable brevemente en español sobre su rutina diaria. Voy a grabar su voz con una grabadora audio digital en HU 286, el cual es un salón que voy a reservar para este estudio. También, usted va a completar un cuestionario sobre su experiencia con el español. La grabación y el cuestionario van a durar no más de media hora.

RIESGOS:
Es posible que usted sienta vergüenza al hacer las pruebas o al ser grabado/a. También, existe la posibilidad que la grabación dure más que el tiempo esperado si ocurren dificultades técnicas.

BENEFICIOS:
Usted no va a recibir ningún beneficio directo por participar en esta investigación. Usted no va a recibir crédito extra en ninguna clase que se ofrece en University at Albany por su participación en esta investigación. Sin embargo, su participación puede ser útil para investigaciones futuras dentro del campo de la adquisición de segundas lenguas.

CONFIDENCIALIDAD:
Toda la información confidencial y nadie, menos que yo, va a tener acceso a ella. Voy a ser la única persona en posesión de su grabación, y voy a borrar su grabación de mi grabadora digital y mi computadora después de la conclusión de la investigación. Todos los datos electrónicos se van a guardar en mi grabadora digital o en mi computadora personal. Este documento y el cuestionario se van a archivar seguramente en un archivador en mi oficina (HU 282). Aunque soy la única persona con acceso a su grabación, es posible que le pida a mi profesor que escuche su grabación para asistirme con el análisis lingüístico. Usted también puede escuchar su propia grabación. Su nombre solo se requiere en este documento si y cuando usted decida participar en este estudio. No le voy a pedir su nombre después de empezar a grabar. Su cuestionario se conecta a su grabación por un número aleatorio que le voy a asignar inmediatamente antes de empezar a grabar. Su nombre no va a aparecer en ninguna grabación ni en los
archivos electrónicos. Si quiere descontinuar su participación en cualquier momento del estudio, sus datos serán destruidos.

PARTICIPACIÓN Y DESCONTINUACIÓN:
Su participación en esta investigación es completamente voluntaria. Aun después de dar consentimiento de participar, usted puede descontinuar su participación en el estudio sin ninguna consecuencia y sus datos serán destruidos. Es posible que yo pudiera terminar su participación en esta investigación sin su consentimiento en las siguientes circunstancias: usted es hablante nativo de un idioma además de inglés o español; usted ha proporcionado información falsa/incorrecta en su cuestionario y usted no está disponible para hacer correcciones; es difícil o imposible usar su grabación por dificultades técnicas y usted no está disponible para repetir la grabación. Es posible que ocurran otras dificultades impredecibles que resulten en la terminación de su participación sin su consentimiento.

IDENTIFICACIÓN DE LA INVESTIGADORA:
Si usted tiene preguntas o preocupaciones, puede contactarse conmigo, Alice Krause:
Número de teléfono: 518-442-4112
Correo electrónico: ak4168@albany.edu

O puede contactarse con mi profesor/supervisor, Maurice Westmoreland:
Número de teléfono: 518-442-4246
Correo electrónico: mw908@albany.edu

DERECHOS DEL PARTICIPANTE:
Puede retirar su consentimiento y descontinuar su participación en este estudio en cualquier momento. Una copia de este documento se queda con los archivos de esta investigación. También, usted va a recibir una copia.
Si tiene preguntas o preocupaciones sobre sus derechos como participante en esta investigación, las cuales la investigadora no ha contestado, puede contactar Albany Office of Regulatory Research Compliance: 518-442-9050 o orrc@uamail.albany.edu.

Por favor, responda a la siguiente pregunta antes de firmar:
¿Acepta ser audio-grabado/a? Circule:  SÍ  NO

Firma del participante
He leído este documento y doy mi consentimiento para participar en esta investigación.

Nombre __________________________________________________________

Firma __________________________________________________________

Correo electrónico _________________________________________________

Fecha ____________________
Appendix C

Background Questionnaire

My name is Alice Krause, and I am doing research for my dissertation. Please fill out the following questionnaire to give me an idea of your background in Spanish. All information is confidential. Thank you.

1) Native language(s) ________________________________

2) Age ____________________

3) What Spanish courses have you taken or are you currently enrolled in at the University at Albany? (check all that apply)

ASPN 100 _______ Semester and year ____________________________
ASPN 101 _______ Semester and year ____________________________
ASPN 103 _______ Semester and year ____________________________
ASPN 104 _______ Semester and year ____________________________
ASPN 206 _______ Semester and year ____________________________
Other(s) __________________________________________________________

4) Have you taken Spanish classes at another university? Yes _____ No _____
Please elaborate as to what level of Spanish that you took at another university or if you hold a degree in Spanish from another university.
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

5) Did you take Spanish in High School? Yes _____ No _____
How much? _____________________________

6) How long (total) have you been studying Spanish?
________________________________________

7) Other languages studied in college or high school?
________________________________________ When? _______________ How long? _______________
________________________________________ When? _______________ How long? _______________
________________________________________ When? _______________ How long? _______________
8) What language(s) are spoken at home?

__________________________________________________________________

9) Have you ever traveled to a Spanish-speaking country? Yes _____  No _____

<table>
<thead>
<tr>
<th>Where</th>
<th>When</th>
<th>Duration of stay</th>
<th>Did you speak Spanish there?</th>
</tr>
</thead>
<tbody>
<tr>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>______</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
</tbody>
</table>

10) Do you speak or hear Spanish in any place other than home or school?
Yes _____  No _____

Where? __________________________________________________________

**Participant number** (to be filled in by researcher)________________________
Appendix D

Cuestionario – Hablantes nativos de español
Me llamo Alice Krause y hago investigación para mi tesis doctoral. Por favor, complete este cuestionario sobre su experiencia con el español. La información es confidencial.
Muchas gracias.
1) Idioma(s) nativo(s) ________________________________
2) Edad __________________________
3) País de origen ________________________________
4) ¿Recibió su educación en un país hispanohablante?
   Escuela primaria Sí No ¿Dónde? __________________________
   Escuela secundaria Sí No ¿Dónde? __________________________
   Universidad
   Subgraduado Sí No ¿Dónde? __________________________
   Graduado Sí No ¿Dónde? __________________________
5) ¿Estudia o ha estudiado otras lenguas?
   Idioma __________________________ Duración de estudio __________________________
   Idioma __________________________ Duración de estudio __________________________
   Idioma __________________________ Duración de estudio __________________________
   Idioma __________________________ Duración de estudio __________________________
6) ¿En qué países ha vivido?
   País __________________________ Duración de estadía __________________________
   País __________________________ Duración de estadía __________________________
   País __________________________ Duración de estadía __________________________
   País __________________________ Duración de estadía __________________________
7) ¿Qué lengua(s) se habla(n) en casa?

Número de participante (para el uso de la investigadora)

_____________________________