**Modern and Historical Factors affect Spanish vowel duration: a corpus study of lexical frequency and root words**

Kyle Parrish

Department of Spanish and Portuguese, Rutgers University

Spanish 511

Dr. [Thomas Stephens](mailto:tstephens@spanport.rutgers.edu)

December 9, 2020

**Abstract**

The present study will aim to investigate Spanish vowel duration for two main reasons. Firstly, it will, in Spanish, replicate recent studies in English which found that higher lexical frequency (i.e. a greater number of occurrences in a corpus of natural speech) resulted in shorter durations of vowels. Secondly, it will consider an additional variable that has not been considered in recent studies involving vowel duration; historical vowel duration (whether a vowel was long or short in Classical Latin). In order to examine these variables, two experiments will be carried out. First, a corpus study of spontaneous spoken Spanish will be conducted in which words will be extracted, segmented and analyzed. These words will be normalized in duration by taking into account speaking rate, word predictability and syntactic category. Then, the relationship between these normalized durations will be compared to their frequency (number of occurrences) in the corpus. It is expected that higher frequency words will systematically shorten in their duration. In the second experiment, these durations, further normalized for frequency, will be compared to their root words of Classical Latin. If historical durational differences persist in modern Spanish, though no longer phonemic, they should show a significant effect in this analysis. These findings have implications for models of sound representation, such as neo-generative models (Levelt, 1980), exemplar models (Johnson, 1997) and hybrid models (Pierrehumbert, 2016; German et al., 2013).

**Modern and Historical Factors affect Spanish vowel duration: a corpus study of lexical frequency and root words**

Research has found that frequency information is encoded in vowel length in English production, such that higher frequency words shorten relative to low frequency homophone pairs (Gahl, 2008). This finding suggests that the lexical item is the locus of frequency information in a language, rather than an abstract phonological level. However, previous studies have not considered the role of historical vowel length in durational differences in modern vowels, such as the potential effects of the remnants of phonemic durational differences found in Classical Latin. In order to address whether modern Spanish retains different lengths for vowels that were once phonemic in Classical Latin, the current study will investigate whether vowel durations in Spanish are sensitive to similar frequency effects found in English, and whether historical vowel length plays a role in durational differences in the production and perception of words in Spanish. These findings are relevant to the conceptualization of a phoneme, and may inform models of sound representation.

In debates regarding the representation of sound structure in the mind, the factors which affect vowel duration have been at the forefront of much research. On one hand, classical generative theory and neo generative models posit that sounds are represented as individual phonemes (Chomsky & Halle, 1968; Levelt, 1980), and that variation in production such as durational differences are noise ignored by the parser during speech comprehension and production, and that the only information extracted from a speech signal is the identification of a phoneme. On the other, so called exemplar models posit that the phonological system is primarily emergent, or determined by use, and that sound segments encode information such as lexical frequency and indexical information about speakers (Johnson, 1997, 2006; Hintzman, 1986). As a compromise to these two extremes, hybrid models, such as Pierrehumbert’s model (2016) do not deny the existence of abstract categories (phonemes), but also suggest that access to these categories are affected by a myriad of social factors, such that comprehension and production include both a phonological parse of abstract categories, and a detailed phonetic trace of individual differences and social information encoded in individual utterances, in a non-systematic and word-by-word basis. In order to research these opposing theoretical perspectives of sound representation, vowel duration has been examined, since neo generative accounts do not consider durational information to be encoded at the phonemic level. Broadly, studies have found that vowel durations are sensitive to extra-linguistic factors, such as lexical frequency (Gahl, 2008). Differences in the frequency of words have been shown to affect the productions of vowel sounds, including their duration. For instance, in a corpus study of English homophone pairs, Gahl found that, when key variables, such as speaking rate, syntactic category, and word predictability are controlled, vowel durations of more frequent content words are shorter than vowel durations of less frequent content words (2008). This finding provides evidence for both exemplar models and hybrid models, since homophones are thought to be made up of the same phonemes, yet their duration was modulated by frequency effects. It is unclear if frequency information affects vowel duration in Spanish, since, to the author’s knowledge, there have not been similar studies to Gahl (2008) done in Spanish.

Naturally, for a model to make claims on the nature of the representation of sound structure in the human mind, their predictions must hold up cross-linguistically, since languages change in distinct ways. For instance, Spanish, derived from Latin, once contained phonemic duration. If listeners tune in to fine phonetic detail in listening, it is possible that word-by-word differences may persist between modern Spanish words which derive from Classical Latin words with long versus short vowel sounds, even when they are no longer phonemic. Should such a difference be found, it would support exemplar models, since word-by-word detailed acoustic information would be both perceived and stored for later production by speakers of modern Spanish. Again, to the author’s knowledge, the present study will be the first to consider a role for the vowel duration of root words on modern vowel durations in conjunction with other factors reported in the literature, such as lexical frequency.

In order to both confirm that lexical frequency modulates vowel length in Spanish and to investigate whether historical vowel length persists in modern Spanish, this study will conduct two analyses of vowel durations in corpus data of spoken Spanish (the C-ORAL-ROM corpus, Cresti et al., 2002). Vitally and in line with previous studies, the present study will statistically control for speaking rate, word predictability and syntactic category and conduct 2 linear regression analyses. The first will examine vowel duration of stressed syllables in order to investigate whether frequency information is encoded in vowel duration in Spanish. The second analysis will, after normalizing duration for frequency, consider the vowel duration of root words relative to derived words in modern Spanish in order to provide evidence that historical linguistic information, such as vowel duration, may still play a role in the realization of Spanish words, despite the loss of phonemic duration in modern Spanish.

**Literature Review**

The body of this paper will begin by briefly describing the models of sound representation, focusing on Pierrehumbert’s hybrid model (2016, German et al., 2013) in comparison to both exemplar models (Johnson 1996; Hintzman, 1986) and neo-generative models of phonological representation (Levelt, 1980). Following this brief summary of the theoretical background, the historical development of modern Spanish vowels and their loss of phonemic duration will be covered. Finally, factors which affect vowel length cross-linguistically will be overviewed with a particular focus on lexical frequency. It will be argued that language cannot be truly and thoroughly analyzed by solely considering its modern usage and status, and that historical features of language, such as phonemic duration, may provide insight in greater debates, such as the debates regarding sound structure in the mind. Additionally, the present study will provide cross-linguistic evidence that word frequency information is encoded in vowels by their differences in duration by examining Spanish language.

***Models of sound structure***

Broadly, the models of sound representation fall under two main categories. Firstly, abstractionist models, also referred to as generativist or neo-generativist models, posit the existence of an independent and discreet phonological system. On the other hand, exemplar models (also sometimes referred to as episodic models) suggest that abstraction is a by-product of the variation encountered during language use, and that these exemplars guide perception and production, rather than independent phonological representations. In response to the predictions of these models, Pierrehumbert’s hybrid model aims support the existence of a phonological abstraction in comprehension and production, but also that detailed phonetic traces of episodes which affect access to representations.

**Classical phonemes and the neo-generative architecture.** The classical notion of the phoneme was put forward in the work *The Sound Pattern of English* (Chomsky & Halle, 1968), and was further developed by Levelt (1980), in what is referred to as neo-generative architecture of sound structure. In this view, phonemes are categorical abstractions which serve as the basis for comprehension and production. Crucially, acoustic variation in the production of phonemes in this view is seen as mere noise. This noise is thought to be largely ignored by the parser during comprehension, and that only the identification of a phoneme is extracted from the speech signal.

As a result, neo-generative accounts of sound representation cannot account for fine-tuned phonetic detail, such as duration, being systematically produced that has been found in recent studies. For instance, Gahl (2008) examined word durations of homophone pairs in English words in the Switchboard Corpus (Deshmukh et al. 1998; Godfrey et al., 1992) and found that higher frequency homophone pairs had significantly shorter durations than their low frequency counterparts. This finding, which supports the idea that word frequency information is encoded in vowel duration, is not a factor that can purely be explained by the use of abstract representations only during speech production. Since the Gahl study investigates homophone pairs, the same phonological representations would be assessed prior to speech production. If this were the case, then there should have been no difference in high frequency and low frequency homophone pairs. Of course, this was not the case. High frequency words shortened, despite being made up of what neo-generative models would consider to be the same phonemes. This result suggests that frequency information cannot be mere noise in the acoustic signal, given its systematic effect on vowel duration, and that it must be attended to and stored in memory by the listener.

**Exemplar models.** As opposed to neo-generative abstractionist models, exemplar models (Johnson 1996; Hintzman, 1986) predict that experience based exemplars, rather than phonemes, serve as the basis for comprehension and production. These exemplars are detailed acoustic signals stored in memory, that also include indexical information about the speaker, such as speaker gender or age. In this view, factors such as recency effects and durational differences can be explained, since production and perception is guided by the exemplars stored in memory that are primarily acoustically based. As opposed to neo-generative accounts, detailed information is attended to by the listener and stored for later use.

**Hybrid models.** As a compromise to exemplar and neo-generative models, hybrid models have been formulated, such as Pierrehumbert’s model (2016). This model is consistent with neo-generative models regarding the existence of the phoneme as an abstract category necessary for comprehension and production, but that access to this representation may be affected by the stored phonetic details of utterances in memory (the exemplars). Unlike neo-generative models, this fine acoustic detail is not seen as mere noise that is not attended to nor stored by the listener. Evidence for Pierrehumbert’s model can be seen in systematic learning across words, such as the retuning of a phonemic category based on recency effects. For instance, Mcqueen et al. (2006) presented Dutch speakers with words they knew ending in a fricative [f] or [s]. The authors created a fricative continuum form [f] to [s] and found an ambiguous point in the middle which they referred to as [?], such that [?] could equally be considered to be [f] or [s]. During a training phase, participants heard Dutch words which they knew to end in [s], but were replaced with [?], while another group heard words they knew to end in [f] which was replaced with [?]. The results revealed that, when hearing new words ending in [?], the participants’ training group modulate their categorization of [?]. That is, if participants heard [?] in place of [s] in words that they knew to end in [s], they generalized this trend to new words. Neither a purely abstract category that does not incorporate recency effects, such as in the neo-generative accounts, nor a word specific acoustic and purely exemplar based account can explain these results. Pierrehumbert’s model, however, suggests that rapid retuning of categories may occur based on recent exposure to exemplars. In other words, phonological representations are not impervious to change based on social factors such as recency of use or retuning, and, exemplars alone cannot explain the generalization across words.

***Historical vowel duration: from Latin to Spanish***

Spanish is a Romance language which was derived from a spoken form of Latin (referred to as Vulgar Latin) (Resnick & Hammond, 2011). In turn, Classical Latin existed at the same time and served as the basis for Vulgar Latin. The vowel system of Classical Latin was made up of 10 vowels which consisted of 5 different vowel qualities of either long or short duration. Importantly, in Classical Latin, vowel duration was phonemic (Resnick & Hammond, 2011), meaning that minimal pairs could be observed simply by changing the vowel duration. However, durational distinction began to become lost in Vulgar Latin, such that the vowel system was reduced to 7, with /o/ and /e/ having open and closed phonemic variants, which would lead to the diphthongs [ue] and [ie] in modern Spanish. As a result, vowel duration is no longer phonemic in Spanish. However, previous studies suggest that listeners attend to acoustic cues beyond phoneme identification, including durational differences (Gahl, 2008). Despite the lack of phonemic contrast in modern Spanish vowels, it is possible that word-by-word differences persist in modern Spanish words which derive from root words with long vowel sounds versus those that derive from words containing short vowels.

***Factors which affect vowel length in English and Spanish***

A myriad of factors have been demonstrated empirically to affect vowel durations in a systematic fashion, although most studies have focused on English. Most relevant to the present study, lexical frequency has been demonstrated to shorten frequent words relative to less frequent words in English (Gahl, 2008). In study of the Switchboard Corpus (Deshmukh et al. 1998; Godfrey et al., 1992). Gahl (2008) showed that homophone pairs in English, such as *time* and *thyme* were produced differently depending upon the lexical frequency of the member of the homophone pair. Frequency was determined by the number of occurrences of a word within the Switchboard corpus. High frequency words, like *time* tended to shorten, whereas low frequency words, like *thyme*, systematically were produced with longer vowel durations. This finding was significant because it suggests that the word is the locus of frequency information, rather than a purely abstract category, such as the phoneme.

In addition to frequency, factors such as local speaking rate, predictability from neighboring words, position relative to pauses, syntactic category, and orthographic regularity also have been found to affect word durations (see Gahl, 2008 for a review). In Spanish, syllable onsets affect vowel duration; simple onset (one segment) results in longer vowel duration, whereas complex onsets (two segments) results in vowel compression (i.e. a shorter duration) (Aldrich & Simonet, 2019).

The study of Spanish vowel duration allows for yet another variable to be examined; historical duration. Given that many studies have focused on English, the study of Spanish allows for both the cross-linguistic testing of frequency effects found in English, and whether historical durational differences which cease to be phonemic still convey systematic differences in modern Spanish. Taken together, both frequency and historical effect on duration may provide a more robust understanding of the nature of phonological representations, since frequency modulated duration may be systematic and historical duration effects may be word-by-word. Should historical durations be correlated with modern durations in Spanish vowels, this would suggest that non-phonemic variation in realizations of Spanish words are perceived, stored and produced by speakers, and would suggest that there are some word specific effects, in addition to lexical frequency, which affect duration.

***The present study***

The present study will aim to inform Pierrehumbert’s hybrid model of sound structure (Pierrehumbert, 2016) by examining the duration of vowel in spontaneous spoken Spanish and determining (1) whether vowel duration is systematically modulated by lexical frequency and (2) whether word-by-word durational differences in Spanish vowels are correlated with root word durations. The data will be sourced from the C-ORAL-ROM corpus, a multilingual corpus of spontaneous speech containing around 1.2 millions words, including Spanish (Cresti et al., 2002), and vowels will be segmented and measured for duration in PRAAT (Boersma & Weenink, 2018). The present study is guided by the following research questions:

RQ1

Under statistical control for known variables impacting vowel duration, is vowel length in Spanish modulated by lexical frequency?

Based on the findings of a similar study in English (Gahl, 2008), it is predicted that frequency information will be encoded in vowels. Just as was the case with English vowels, it is predicted that Spanish vowels will shorten in high frequency words relative to low frequency words. Duration of these vowels will be normalized based on syntactic category, speaking rate, and word predictability.

RQ2

When frequency is controlled for, do historical vowel duration differences persist in modern Spanish?

Although vowel duration in modern Spanish is not phonemic, durational differences in words do seem to be attended to by listeners and duration has been shown to encode frequency information in production. Based on these findings, it is possible that durational difference may be present at a word by word level. Should vowel duration in modern Spanish be correlated to root word length, it would support exemplar models and word by word effects of duration in non-phonemic contexts over a significant period of time and throughout the generations that have seen Vulgar Latin develop into modern Spanish.

**Methods**

**Materials and procedure**

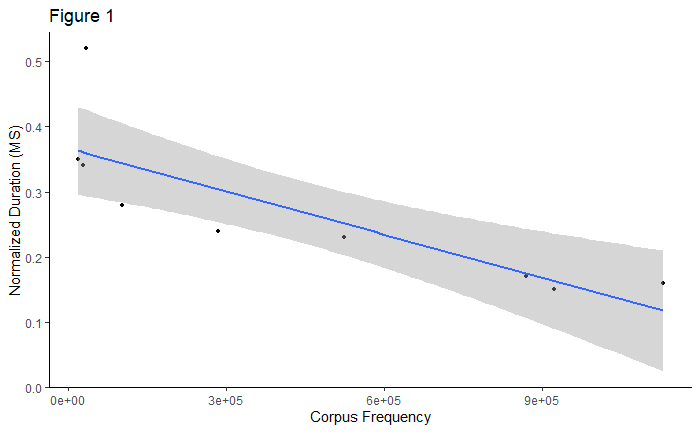
Stimuli will be gathered from the C-ORAL-ROM corpus, a multilingual corpus of spontaneous speech containing around 1.2 millions words (Cresti et al., 2002). The words will be segmented into words and vowels in PRAAT (Boersma & Weenink, 2018) and measured for duration. Additionally, words will be measured within an utterance per speaking rate by using syllables per second and durations will be normalized. Once words are segmented and labeled for vowels, the number of occurrences in the corpus will be recorded as a measure of lexical frequency. In addition to frequency, syntactic categories and word predictability will be statistically controlled for, since predictable words have been shown to have shorter durations (Gahl, 2008). Predictability will be measured by using bigram frequency, as has been done in previous studies (Gahl, 2008). Bigram frequency refers to the likelihood of a word occurring based on neighboring words. A higher bigram frequency means that a particular word is more likely to occur when preceded by another word in particular, and is a measure of predictability. It is expected that a word with a low bigram frequency will have a longer duration, since it is not as predictable as a word with higher bigram frequency.

**Analyses**

The present study will conduct two analyses of Spanish words from the corpus. The first will examine vowels across all words and normalize their durations based on syntactic category, word predictability and speaking rate and analyze these normalized durations against word frequency (measured by occurrences in the corpus). Should frequency affect vowel length, then vowels should become shorter as frequency increases. The second analyses will examine historical vowel length in relation to modern duration when frequency is controlled for. Following the first analysis, the second will investigate the variability found in durations by comparing whether the root word, when the root word is Latin, had a long or short vowel in Classical Latin. Root words will be sourced from the etymological dictionary by Corominas (Corominas, 1961). A linear regression will be run to determine whether the relationship between historical vowel length (code as either long or short, 1 or 0), has a significant relationship with modern vowel length when frequency is controlled.

**Predicted Results**

If frequency information is contained in the vowels of modern Spanish words, then duration should decrease as frequency increases. In a pilot study, a Spanish professor was recorded during a graduate Spanish course at a university in the northeast United States. His productions were spontaneous and recorded for purposes of the course. Utterances containing *-ar* verbs in Spanish were taken and segmented in PRAAT. Durations were normalized by controlled for speaking rate (syllables per second) and plotted against frequency (from the Corpus del Español: NOW). Following the plot, a linear regression was done in R (Wickham, 2017) to produce a line of best fit (figure 1). The preliminary results suggest that frequency information is encoded in Spanish vowels, similar to findings in English vowels (Gahl, 2008).



Subsequent analyses of the effects of lexical frequency on vowel duration will involve an analysis of a broader class of words and durations will be normalized. Additionally, all five vowel sounds of Spanish will be analyzed and compared, in order to determine whether there are significant differences per vowel.

**Conclusion**

The goals of the present study are twofold. Firstly, the present study will aim to provide evidence that frequency information is encoded in Spanish vowels, such that shorter durations are indicative of more frequent words. Secondly, the present study presents a novel variable in word duration studies, historical vowel duration. The results of both goals have implications for models of sound representation, such as the neo-generative models (Levelt, 1980), exemplar models (Johnson, 1997) and hybrid models (Pierrehumbert, 2016), since these findings suggest that durational differences are perceived and produced by listeners, even when they are not phonemic. If frequency information is found to affect word duration, it would best be explained by a hybrid model of sound representation, such as Pierrehumbert’s model (2016), since frequency information is both systematically implemented and subject to word-by-word effects. However, should the second experiment reveal that Spanish vowel length is correlated to Classical Latin vowel length, then both exemplar models and Pierrehumbert’s model could explain these results, since this would be indicative of word-by-word non-systematic effects. Additionally, neo-generative models could only explain the results if no relationship exists between frequency, historical vowel duration and modern duration. That is, if vowel duration is consistently produced differently when it is no longer phonemic, this would be evidence against the neo-generative architecture, since non-phonemic acoustic detail is not being ignored by speakers as noise.

**References**

Aldrich, A. C., & Simonet, M. (2019). Duration of syllable nuclei in Spanish, *Studies in Hispanic and Lusophone Linguistics, 12*(2), 247-280. doi: <https://doi.org/10.1515/shll-2019-2012>

Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [computer program]. Retrieved from <http://www.praat.org/>

Corominas, J. (1961). *Breve diccionario etimológico de la lengua castellana.* Editorial Gredos.

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

Cresti, E., Moneglia, M., Bacelar F., Moreno A., Veronis J., Martin P., Choukri K., Mapelli V., Falavigna D., Cid A., & Blum C. (2002). *The C-ORAL-ROM Project. New methods for spoken language archives in a multilingual romance corpus* In Procs. of LREC-2002 pp. 2-9.

Deshmukh, N., Ganapathiraju, A., Gleeson, A., Hamaker J., & Joseph Picone. (1998). Resegmentation of Switchboard. Paper presented at the International Conference on Spoken Language Processing, Sydney, Australia

Gahl, S. (2008). “Time” and “Thyme” Are Not Homophones: The Effect of Lemma Frequency on Word Durations in Spontaneous Speech. *Language (Baltimore), 84*(3), 474–496. <https://doi.org/10.1353/lan.0.0035>

German, J. S., Carlson, K., & Pierrehumbert, J. B. (2013). Reassignment of consonant allophones in rapid dialect acquisition. *Journal of Phonetics, 41*(3-4), 228–248. <https://doi.org/10.1016/j.wocn.2013.03.001>

Godfrey, J.J. , Holliman, E. C, & McDaniel, J. (1992). SWITCHBOARD: telephone speech corpus for research and development. *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing, 1*, 517–520 vol.1. <https://doi.org/10.1109/ICASSP.1992.225858>

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411-428.

Johnson, K. (1997). Speech perception without speaker normalization. In K. Johnson & J. W. Mullennix (Eds.), Talker variability in speech processing (pp. 145-166). San Diego: Academic Press. Johnson, K. (2006).

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics, 34*, 485-499.

Levelt, W. J. M. (1980). *Speaking*. Cambridge, MA: MIT Press.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological Abstraction in the Mental Lexicon. *Cognitive Science, 30*(6), 1113–1126. https://doi.org/10.1207/s15516709cog0000\_79

Pierrehumbert, J. B. (2016). Phonological Representation: Beyond Abstract Versus Episodic. Annual Review of Linguistics, 2(1), 33–52. https://doi.org/10.1146/annurev-linguistics-030514-125050

Resnick, M., & Hammond, R. (2011). *Introducción a la historia de la lengua española* (2a ed.). Georgetown University Press.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. New York,

NY: Springer-Verlag.

Wickham, H. (2017). tidyverse: Easily Install and Load the ‘Tidyverse.’ R

package version 1.2.1. Available online at: https://CRAN.R-project.org/

package=tidyverse (accessed February 24, 2019)