

# The Production and Perception of Vowel Height and Duration

by

Adam Stone

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in  
partial fulfillment of the requirements for the degree of

Master

in

Cognitive Science

Carleton University  
Ottawa, Ontario

© 2015

Adam Stone

## Abstract

High vowels are generally shorter than low vowels: there is a positive correlation between F1 and duration in English and cross-linguistically (Heffner 1937; Elert 1964; Äimä, 1918). This paper argues that the cross-linguistic height/duration correlation might be explained perceptually: high vowels inherently *sound* shorter than low vowels. Study 1 analyzed Chilean Spanish vowels to determine whether this correlation is physiological in nature or the result of linguistic rules, finding that the correlation is linguistically-specified. To account for the cross-linguistic occurrence of the correlation, Study 2 tested if speakers perceive shorter high vowels in a forced-choice perception task. Results indicate that high vowels indeed sound shorter, and that this vowel categorization ability is partially learned.

## Acknowledgments

This is probably the part of my thesis which requires the most careful thought and deliberation. First and foremost, I would like to thank my thesis supervisor and friend Dr. Ida Toivonen, who spent countless hours coaching me and pointing me in the right direction. She also found me a job, directed me to the program itself, involved me in research projects, sent me to countless conferences, introduced me to all the right professors, enabled me to explore my interests in Sámi linguistics, helped me get scholarships, found some amazing books for me, and told me some really great jokes that won't be repeated here. The list goes on. I am unable to thank her enough for the contributions she has made to my academic life.

I would also like to thank Alex Olsen, who gave me a very good reason to visit Chile. Not only was she a great guide across the country, but assisted in the translation of materials from and into Spanish, found participants, provided lodging and transportation, and gave invaluable moral support throughout the whole process of both the Chilean Spanish experiment and my degree.

I would like to thank my father Jim for being an unfathomably deep well of knowledge and experience who would always tell me what to expect, as he's done "this sort of thesis thing" before. On this note, my brother Ben was helpful in giving me the downtime I needed to chill out from writing, and my mother Michaela gave me unmeasurable moral support.

Last but certainly not least, I would like to thank my LLI lab for first of all, existing, and second of all, for containing some of the most intelligent and warm-hearted people I have ever met. Each one of them have left a lasting impression on me, and have shared my triumphs and frustrations. I would also like to thank my Thesis Defense committee (Drs. Ida Toivonen, Lev Blumenfeld, John Logan, Beth MacLeod, and Kasia Muldner) for being tough, fair, and kind.

## Table of Contents

1. Introduction .....	9
1.1 Evidence for DHC .....	9
1.1.1 Duration as a Primary Cue .....	10
1.1.2 Duration as a Secondary Cue .....	11
1.1.3 Duration across Languages .....	12
1.2 Explanations for the Duration Height Correlation .....	13
1.2.1 The Mechanical/ Phonetic Hypothesis .....	13
1.2.2 The Controlled/ Phonological Hypothesis .....	14
1.2.3 Integration of Mechanical and Controlled hypotheses .....	15
1.2.4 Cognitive Evidence .....	16
1.2.5 Neurological Evidence .....	16
1.3 Duration Perception .....	17
1.4 Testing DHC .....	19
1.4.1 Studies 1 and 2 .....	19
2. Study 1: Production .....	20
2.1 Procedure .....	20
2.2 Participants .....	21
2.3 Measures .....	21
2.3.1 Between-Category Measurements .....	21
2.3.2 Within-Category Measurements .....	22
2.3.3 Predictions .....	22
2.4 Results .....	23
2.4.1 Volunteered subjective reports .....	23
2.4.2 Between-Categories .....	24
2.4.3 Within-Categories .....	25
2.5 Discussion .....	27
3. Study 2: Perception .....	28
3.1 Procedure .....	29
3.2 Participants .....	30
3.3 Measures .....	31
3.3.1 Choice Task .....	31

3.3.2	Response Time.....	31
3.3.3	Inter-Stimulus Intervals .....	32
3.4	Materials.....	33
3.5	Design.....	34
3.6	Predictions.....	36
3.6.1	Vowel Pairs.....	36
3.6.2	Natural vs. Modified Duration.....	38
3.6.3	Real vs. Nonsense Words .....	38
3.6.4	Short vs. Long ISI.....	39
3.7	Results.....	39
3.7.1	Subjective Reports .....	39
3.7.2	Choice Data.....	41
3.7.3	Response Time Data.....	49
3.8	Discussion .....	54
3.8.1	Overall Trends .....	55
3.8.2	Vowel Pairs.....	56
3.8.3	Natural vs. Modified Duration.....	58
3.8.4	Real vs. Nonsense Words .....	60
3.8.5	Short vs. Long ISI.....	62
3.8.6	Limitations .....	64
4.0	General Discussion .....	66
4.1	Necessary Future Research .....	68
5.0	Conclusion .....	70
6.0	<i>References</i> .....	72

## List of Tables

Table 1: Mean stimuli results for f1 height and f2 backness and duration. ....	33
Table 2: Experimental Conditions for the English stimuli .....	35
Table 3: Experimental Conditions for the Swedish stimuli .....	35
Table 4: Summary of within-pair comparisons for English vowels .....	43
Table 5: Summary of within-pair comparisons for Swedish vowels .....	44
Table 6: English length choices between natural and modified duration conditions .....	45
Table 7: Swedish length choices between natural and modified duration conditions .....	46
Table 8: English real vs. nonce vowel choices .....	47
Table 9: English Choices for Long and Short ISI.....	48
Table 10: Swedish Choices for Long and Short ISI .....	49
Table 11: English RT Results – Repeated Measures ANOVA.....	51
Table 12: Swedish RT Results – Repeated Measures ANOVA .....	54

## List of Figures

Figure 1: Representation of how languages use contrastive duration as a spectrum.....	12
Figure 2: Participants' Average Vowel Durations.....	24
Figure 3: Participants' Average f1 Height.....	25
Figure 4: Within and between-category results for Study 1 .....	26
Figure 5: Main stimuli presented to participants .....	30
Figure 6: English vs. Swedish Length Choices .....	42
Figure 7: RT for nonce/ real words across ISI.....	52
Figure 8: Vowel pair RT across duration conditions .....	53
Figure 9: Response Time and Length Choice for English Vowels.....	57
Figure 10: English and Swedish Length Choices, RT and ISI .....	63

## List of Appendices

Appendix A.....	83
Appendix B.....	85
Appendix C.....	87
Appendix D.....	88
Appendix E.....	90
Appendix F.....	99
Appendix G.....	103
Appendix H.....	104
Appendix I.....	108



## 1. Introduction

High vowels are generally shorter than low vowels cross-linguistically. There are two competing explanations why this trend, known henceforth as the Duration-Height Correlation (DHC), exists. The mechanical hypothesis appeals to physiology, whereby low vowels take longer to pronounce because they involve a greater degree of jaw opening. The controlled hypothesis states that duration targets are specified by a language's grammar, and are thus speaker-controlled. Though the mechanical hypothesis explains the cross-linguistic nature of this correlation, it does not account for the findings discussed by Solé & Ohala (2010) in section 1.2.1. Inversely, the controlled hypothesis accounts for these findings, but not the cross-linguistic tendency. Study 1 investigates whether DHC is mechanical or controlled by measuring Chilean Spanish vowels, and finds that DHC is controlled. To explain the cross-linguistic nature of the controlled DHC, Study 2 hypothesizes that speakers perceive high vowels as shorter, and investigates the extent to which DHC is learned. Together, this thesis explores both how speakers *produce* vowels and how they *perceive* them.

### 1.1 Evidence for DHC

Evidence from languages such as English (Heffner, 1937; House & Fairbanks, 1953; Peterson & Lehiste, 1960; Scharf, 1962), Swedish (Elert, 1964), Inari Saami (Äimä, 1918), Thai (Abramson, 1962), and Spanish (Navarro-Tomás, 1916) suggests that DHC is common across both languages and language families, and is found in both tonal and non-tonal languages (Järvikivi, Vainio & Aalto, 2010). According to Lehiste (1970, p.18), typical durational differences between high and low vowels should be above the threshold for auditory

discrimination, meaning that they are perceivable by humans and likely serve a linguistic purpose.

Though no reports have found instances where high vowels are generally longer than low vowels in a language, Johnson & Martin (2000) have noted an exception where word-initial high vowels in the Creek language tend to be longer than word-initial low vowels. Such examples are however exceedingly rare.

### 1.1.1 Duration as a Primary Cue

Vowel duration is distinctive in a number of languages, including Finnish and Estonian (Ylinen, Shestakova, Alku & Huotilainen, 2005), Japanese and Norwegian, (Kingston, Kawahara, Chambless, Mash & Brenner-Alsop, 2009), and Swedish and Inari Saami (Bye, Sagulin & Toivonen, 2009). These languages are known to employ *contrastive duration*, where durational changes are a primary cue, and are known as *quantity languages*. As an example, Finnish speakers will readily distinguish between the *tapan* “I kill” and *tapaan* “I meet,” where a single vowel refers to a short vowel, and a double vowel refers to a long one (Bergmann et al., 2007).

There are some differences among quantity languages in how contrastive duration is realized. Swedish uses duration as a primary cue, but takes other factors such as vowel quality into account, whereas Finnish distinguishes meaning by duration alone (Engstrand & Krull, 1994, p.40). Even though other factors than duration may be peripherally involved in quantity languages, duration is still the primary contrastive factor.

An important element to keep in mind when focusing on vowel length is the environment that the vowels occur in. Solé (2007) mentions that vowels behave similarly to directly following

consonants (e.g. the /i/ in *bean* is more nasal than in *beat*, as the consonant /n/ is nasal<sup>1</sup>). This is because of the time required for one's airstream to switch from leaving the mouth to the nose by raising and lowering the velum. Lehiste (1970, p.20) notes that complex articulations yield longer durations than simple articulations, such that vowels occurring before consonants such as /d/ or /g/ are typically longer than before /b/ as the former consonants require two articulators for production, and not one. Similarly, Fischer-Jørgensen (1964) explains that the main factor in determining duration is the length of articulation needed to pass from the vowel to the consonant position.

Because of this, paying attention to a vowel's linguistic environment is paramount to understanding its form. The environment of vowels being analyzed should be controlled, to remove transitional interference from surrounding consonants.

### 1.1.2 Duration as a Secondary Cue

Durational changes can be found in non-quantity languages as a secondary cue, or a biomechanical by-product of the vowel being manipulated through a primary cue (Solé & Ohala, 2010). An example of this is found in English tense/lax vowel distinctions (/i/ vs. /ɪ/), where lax vowels are primarily differentiated from tense vowels by vowel quality (F1, F2, F3 frequencies), with duration playing a secondary role (Kondaurova & Francis, 2008, p.3959).

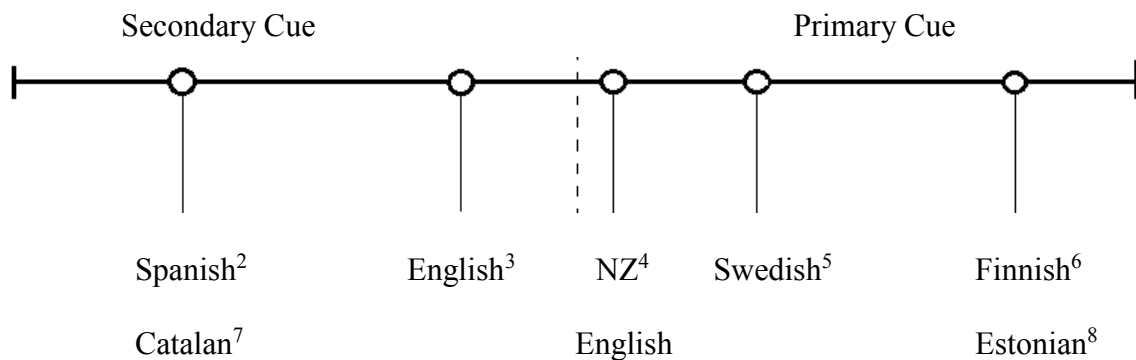
---

<sup>1</sup> Nasal Airstream: Majority of air is blown out through the nose, as the velum (soft palate) closes off the airstream from the mouth (Bergmann et al., 2007)

### 1.1.3 Duration across Languages

To illustrate how contrastive duration occurs in different languages, languages can be represented on a spectrum: on one end, meaning is distinguished entirely by changing the length of an articulation, and on the other, changing a vowel's length makes no phonological distinction. This representation is seen below, in Figure 1.

Figure 1: Representation of how languages use contrastive duration as a spectrum



DHC was found in McAllister, Flege, and Piske (2002) to occur similarly for both English (where duration is a secondary cue) and Swedish (where it is a primary cue) meaning that how duration occurs does not affect the presence of DHC itself. Primary cues are speaker controlled, whereas secondary cues need not be (Solé & Ohala, 2010, p.258).

---

<sup>2</sup> Navarro-Tomás (1916)

<sup>3</sup> Heffner (1937) ; House & Fairbanks (1953)

<sup>4</sup> Bauer & Warren (2008)

<sup>5</sup> Elert (1964)

<sup>6</sup> Ylinen et al. (2005)

<sup>7</sup> Solé & Ohala (2010)

<sup>8</sup> Ylinen et al. (2005)

## 1.2 Explanations for the Duration Height Correlation

### 1.2.1 The Mechanical/ Phonetic Hypothesis

A *mechanical* (physiological) explanation appeals to the physical design of the articulatory system. In this more traditional perspective, low vowels take longer to produce than high vowels because of the extra time to open the jaw (House, 1961; Lindblom, 1967; Lehiste, 1970; Catford, 1977). In Lindblom (1967), tongue position in relation to the palate has an effect, where tongue position is correlated with the degree of jaw opening, leading to the hypothesis that jaw opening is correlated with vowel duration (p. 2). This is observed in Catford (1977), where the greater the distance between the vowel and surrounding consonants' places of articulation, the longer the vowel's articulation (p.197). As DHC occurs cross-linguistically, it is seen as a fundamental characteristic of human language (House, 1961), and is beyond speaker control (Lehiste, 1970). The mechanical explanation is also referred to as *phonetic* by researchers such as House (1961) and Lindblom (1967). As the naming suggests DHC is an inseparable component of vowels.

However, there are some issues with this theory. If DHC were entirely physiological, then the length of vowels in relation to their heights would not be affected by speech rate. In a series of experiments testing the effects of speech rate on vowels' heights, Solé and Ohala (2010) found that duration varied for both English and Catalan, but remained unchanged (thus mechanical) for Japanese. Though these mechanical findings for Japanese suggest a physiological basis, the controlled DHC elsewhere and the fact that the languages differ suggest that DHC is controlled in some languages.

In addition, Lisker (1974) suggests that if low vowels are longer because of the increased time for the jaw to close, transition states at the beginning and end of the vowel would also be consistently longer, and the steady state of the vowel would not be. However, it is in fact the steady state and not the on-and-off transitions that are longer.

Furthermore, Toivonen, Blumenfeld, Gormley, Hoiting, Logan, Ramlakhan, and Stone (2014) suggest that if DHC is mechanical, vowel duration would vary according to f1 differences between different utterances of the same vowel. However, their results suggest that this is not the case.

Finally, if DHC was physiological, then when considering dialect variation in vowel height, one would predict that dialects with higher versions of a particular vowel than neighbouring dialects would also have measurably shorter instances of that vowel. However, Tauberer and Evanini (2009) found no such correlation, suggesting that DHC is not mechanical in nature (p.2214).

### 1.2.2 The Controlled/ Phonological Hypothesis

According to the controlled explanation, DHC results from active (albeit unconscious) control of duration by speakers, perhaps attempting to make vowels discernable and ease information processing (Solé & Ohala, 2010). Duration thus becomes a consistent element in a language's sound system, and each vowel is said to have a specified durational target in a given language's grammar (Lisker, 1974; Tauberer & Evanini, 2009).

Though DHC is stated as "controlled," it is not to say that speakers consciously decide to make high vowels shorter than low vowels, or that changing duration according to vowel height becomes conscious at any point. Instead, the term "controlled" specifies that DHC occurs not as

some physiologically determined behaviour, but instead results from a language's rules learned by its speakers. However, this explanation is problematic, as it indicates that every language just happens to choose the same rules that dictate how duration occurs in relation to height. This leaves unexplained why the DHC is potentially universal.

### 1.2.3 Integration of Mechanical and Controlled hypotheses

As neither interpretations completely account for DHC, a combination of the two hypotheses may be necessary (Solé & Ohala, 2010). Solé (2007) argues that language-specific features (such as DHC) are variants of universal features that are specified in motor commands. This physical system may be elaborated on by speakers in order to increase vowels' distinctiveness. This is observed in non-quantity languages, where vowel duration increases differences between phonemes (Meister & Werner, 2009), and is thus important to increasing distinctiveness, but itself is not responsible for changes in meaning. In support, Gottfried and Beddor (1988) speculate that vowel duration is perceptually important for English and languages where phonologically non-distinct length covaries with spectral properties (p.58). Furthermore, they speculate that this perceptual importance may in turn be influenced by the overall prominence of vowel duration in the language.

Explanations outside of linguistics could bridge the conceptual gap between the mechanical and controlled theories with evidence traditionally not included in linguistic analyses. The sections below will outline cognitive and neurological evidence for the importance of duration in language.

#### 1.2.4 Cognitive Evidence

Duration has been shown to increase intelligibility in vowel qualities for adults, and even more so for infants. In Bohn and Polka (2001), German infants relied on durational cues significantly more than adults, and had greater trouble understanding tokens with vowels with altered durations (p.512). Though adults typically develop the ability to recognize other distinguishing cues of vowel quality such as intensity (Lehiste 1970; p.131), pre-linguistic infants appear to rely more on duration. Furthermore, there are certain conditions where durational differences alone can specify vowel contrasts in adulthood, such as in the presence of noise (Gottfried & Beddor, 1988, p. 58).

In Eilers et al. (1984), where English-speaking pre-linguistic infants distinguished three synthetic vowel lengths embedded in minimal pairs (e.g. /mæt/, /mæd/), indicating that infants use vowel duration as a primary cue to distinguishing between word final stops, which when unreleased, tend to be phonologically ambiguous in isolation (Raphael, 1972; p. 1296).

In adding the possibility that DHC is a cognitively innate feature in humans, a collaborative explanation of DHC becomes reasonable assuming the Chomskyan notion that all human languages are variants of an underlying structure governed with universal rules and themes (Chomsky, 1972). Following this model, DHC would have both physiological and cognitive ‘innateness’ expressed in varying degrees according to the linguistic environment and rule structure, along with aspects such as tenseness, rounding, and backness.

#### 1.2.5 Neurological Evidence

Duration and timing are essential to communication in general, and not just to vowels. Calvin (1983) identified an understanding of timing and sequencing as paramount to language



development, as well as memory and cognition. According to Mateer and Kimura (1977; in Calvin, 1983), timing specializations are integral to both phonemic and motor aspects of language. This is supported by MRI data from Schubotz, Friederici, and von Cramon (2000), where neural networks supporting time perception also involve areas responsible for the coordination of oro-facial movements and temporal planning. Brain regions associated with language production and perception are associated with temporal regulation and interpretation, including the peri-Sylvian language cortex (Ojemann & Mateer, 1979), Broca's area (Sahin, Pinker, Cash, Schomer, & Halgren, 2009; Schubotz et al, 2000), and Wernicke's area (Price, Veltman, Ashburner, Josephs, & Friston, 1999).

### 1.3 Duration Perception

Several studies indicate that listeners perceive high vowels as longer than low vowels, with an inverse relationship to the vowels' actual durations. In Wang, Lehiste, Chuang, and Darnovsky (1976), participants judged the comparative durations of synthesized /i/, /e/, and /æ/ vowels, with each vowel manipulated to either have a level, rising, or falling F0.<sup>9</sup> Though no *significant* effect occurred between the durations of /i/ and /æ/, /i/ was found to be perceived as slightly longer than /æ/. Wang et al. (1976) claim the effects exist due to a compensatory strategy to eliminate durational differences between vowels. Effects were shown for both linguistic and non-linguistic stimuli, suggesting that this compensatory strategy is likely not restricted to speech.

---

<sup>9</sup> Though this study focused on F0 and not F1, Hoemeke and Diehl (1994) noted a near perfect inverse relationship between F0 and F1, indicating that F0 studies are still useful as evidence in F1 research.

Gussenhoven and Zhou (2013) hypothesize (based on the results of Gussenhoven, 2007) that the perceived durational values of vowels are inversely related with their actual measurable values. Like Wang et al. (1976), the researchers cite a compensatory strategy and claim a non-linguistic psychoacoustic origin. Brigner's (1988; in Gussenhoven & Zhou, 2013) results support this non-linguistic origin, as participants rated high tones as sounding longer than low tones of the same actual length. In support of the contrastivity hypothesis, Gussenhoven (2007) explains that this effect may be used by speakers who lower vowels to signify a short duration, or raise vowels to signify a long duration (p.22). Thus Gussenhoven's (2007) theory still accounts for an inverse relationship between perceived and measured vowels, but explains that the compensatory strategy is a (potentially prelinguistic) psychoacoustic phenomenon used by speakers to increase spectral distinctiveness.

It is because of these findings that studying speakers' perception of speech sounds is important. Kingston et al. (2009) provide evidence for an Interactive Model for Speech Perception: when a speaker hears a speech sound, their linguistic knowledge feeds back immediately to their judgments of that signal's perceived acoustic properties. In this model, categorization (i.e. determining which vowel is which) is not distinguished from perception, as linguistic knowledge influences both the listener's bias to respond, and the percept that the response is based on.

However, Kingston et al. (2009) continue on to support the notion of Direct Realism, whereby hearing and perceiving are not distinguished from one another. In direct realism, the perception of a sound cannot transform that sound's acoustic properties, as the speaker must be able to properly perceive that sound to successfully understand the communication. Support of this model contradicts the previous findings in Gussenhoven (2007) and Wang et al. (1976), as

these researchers posit that perceptions of speech sounds have been transformed from a more automated hearing stage. Findings from Study 2 may shed some light on this disparity.

#### 1.4 Testing DHC

Neither of the two traditional competing theories of mechanical and controlled vowel duration have fully accounted for why DHC exists. My thesis will thus operate under the assumption that a collaboration of the two theories can account for DHC. As a result, my evidence indicates that DHC presents itself as controlled, but may have some cognitive underpinnings along with a potential set of physiological constraints, which influences how duration as a secondary cue is grammaticalized.

##### 1.4.1 Studies 1 and 2

Two studies will explore how DHC occurs in a language and how speakers develop intuitions about how long vowels should be. The first study is observational in nature, and aims to both record DHC in Chilean Spanish, and test whether DHC is mechanical or controlled. The second study will test participants' intuitions about different vowels' durations, by testing participants' ability to choose between two words of variable length, how long it takes them to choose, and what levels of cognitive processing these choices may occur on. Though the following two studies have similar goals, they differ in their design and populations. Ethics approval was granted by the Board of Ethics at Carleton University on August 8<sup>th</sup>, 2014 through a Minimal-Risk Application, and is valid until May 31, 2015. Full anonymity is guaranteed to participants in both studies.

## 2. Study 1: Production

This study aims to find evidence for DHC in Chilean Spanish by measuring the height and length of speakers' vowel productions. Evidence from this study will contribute to the DHC literature, and provide a base for which research can be conducted in study 2. The Spanish language is of interest as, along with a relatively low vowel inventory (Payne, 2008), durational differences are smaller than in languages such as English (Zimmerman & Sapon, 1957; p.152). As a divergent dialect (Lipski, 1994), vowel duration may show different patterning in Chilean Spanish than in other dialects.

### 2.1 Procedure

This study replicates the basic design from previous studies on DHC by Toivonen et al. (2014). Participants read 10 lists of minimal pairs and triplets in Spanish, found in *Appendix A*. Each list consists of the same 21 words in randomized order. Participants read each word aloud in the carrier phrase “*decir \_\_\_ para mí*” (“to say \_ to me”). Carrier phrases allow words to occur as if spoken in a sentence, and not be subject to the increased stress and the resulting intensity that occur when in isolation (Peterson & Lehiste, 1960). As reported in Lehiste (1970), intensity can directly affect vowels' quality, and should thus be an important factor when measuring productions. Recordings were made using a TASCAM solid state device, which offers studio quality recordings with minimal background noise, with the assistance of a head mounted microphone. Recordings were analyzed for both f1 and duration in the PRAAT acoustic software (Boersma & Weenink, 2014).

Annotations (combinations of the sound file and manually entered lexical information) were collected and analyzed by a script created by Mietta Lennes (2011) and modified by

Toivonen et al. (2014) which returned information on duration, fundamental frequency ( $f_0$ ),  $f_1$ , and  $f_2$ . The data was exported into SPSS statistics software, and graphs used Microsoft Excel. The front vowels /i/, /e/, and /a/ were focused on, as they are all fronted, and are respectively a high, mid, and low vowel.

## 2.2 Participants

6 native speakers of Chilean Spanish participated in Santiago, Chile, during the summer of 2014. Participants came from different strata of Chilean society, and varied socioeconomically from custodian to CEO of a major bank. This is important to obtaining a representative sample of the Chilean population, as there is strong dialectical variation according to socio-economic status (Bernstein, 2013). Speakers were monolingual to remove interference from English or another L2 language. No rewards were offered for participating in the experiment. Contacts between participants and researcher were established by a Carleton Student on exchange in Chile with cultural and linguistic familiarity with Chilean society. Each participant was provided with a form detailing the instructions of the experiment and with a consent form (see Appendix H).

## 2.3 Measures

To test whether DHC is mechanical or controlled in Chilean Spanish, two separate types of measurements were performed, as detailed below.

### 2.3.1 Between-Category Measurements

Between-Category measurements are between different vowels (e.g. how /i/ and /e/ are measurably different). As a more traditional analysis, between-category measurements are used by researchers such as Solé and Ohala (2010) to observe duration and height distinctions

between various vowels. Evidence from this method has been used by Heffner (1937), House and Fairbanks (1953), Peterson and Lehiste (1960) (to name a few), to indicate measurable differences in duration between high and low vowels. If DHC occurs in Chilean Spanish, I would expect to see measurable differences in f1 and duration between /i/, /e/, and /a/. Though this standard approach often does not have a particular name attached to it, I will use the term *between categories*, as used in Toivonen et al. (2014).

### 2.3.2 Within-Category Measurements

As opposed to between-category measurements, *within-category* measurements focus on different utterances of the *same* vowel. This measurement assumes that no two instances of the same vowel are ever measurably the same (Peterson & Barney, 1952, p.175), and was focused on by Swoboda, et al. (1978, p.334). Here, every utterance of /i/ produced by a speaker would be measurably different from one another (i.e. in terms of f1 height), whether produced in quick succession or isolated from different parts of speech. Interestingly, this occurs even if the speaker attempts to control the vowel utterances to be the same. As within-categories effects are outside speaker control, any changes in f1 height that correlate positively with durational changes would point to a mechanical underpinning (Toivonen et al., 2014). In other words, an instance of /i/ that is higher than a previous utterance of /i/ would be expected to also be measurably shorter.

### 2.3.3 Predictions

If the effect is purely mechanical, I would expect to see the same effect within as between categories. In this case, high vowels would be consistently shorter than low vowels, and higher instances of a vowel would be shorter than lower instances of the same vowel. If the effect is

instead speaker-controlled, there would be a significant difference between categories, but not within categories. These assumptions are identical to Toivonen et al. (2014), who performed the same experiment on Swedish and English, both of which employ contrastive duration more than Spanish.

## 2.4 Results

### 2.4.1 Volunteered subjective reports

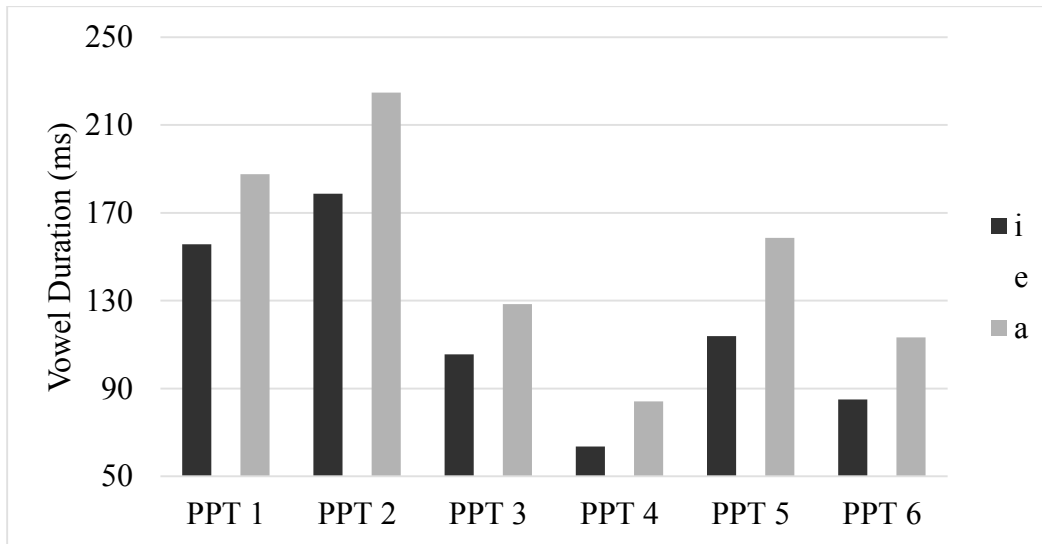
Upon completion, some participants volunteered unsolicited subjective experiences, which proved useful as an informal evaluation of the experimental process. As none of the participants had a scientific background, no comments were offered regarding the experimental process. Participants enjoyed the experiment, and found the instructions and the task simple and straightforward. In confirmation of the guidelines set in the minimal-risk ethics application, no participant reported feeling that their safety, mental/physical health, and/or well-being had been compromised. Though some participants reported the experiment as repetitive, that aspect is neither avoidable nor consequential to the validity of the data.

As the tasks were performed during the day in a standard work week, some participants reported being constrained for time, but eagerly completed the task. Though most were at first puzzled by the underlying reasons for the recordings, they were satisfied in the debriefing portion of the experiment. Aside from an overall faster speech rate from some of the time-constrained participants, recordings were all performed without incident.

## 2.4.2 Between-Categories

Exploratory statistics focused on 1 227 data points. Participants were analyzed separately and then compared to each other as their natural speaking rates differed, as shown in Figure 2.

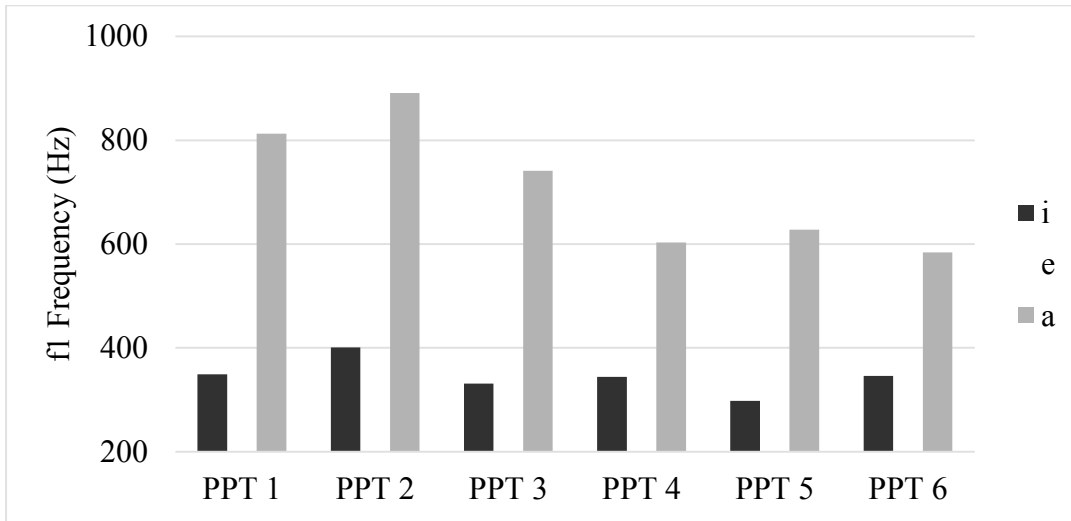
Figure 2: Participants' Average Vowel Durations



Regardless of speech rate, durational differences between vowels form a consistent and expected pattern. This provides support for the hypothesis that DHC is found in Chilean Spanish as it has been found in many other languages. Similar results were found for /l/ in Figure 3.



Figure 3: Participants' Average f1 Height



An independent samples t-test found significant durational differences between /i/ ( $M = 117.39$  ms;  $SD = 44.08$ ) and /e/ ( $M = 132.65$ ;  $SD = 50.13$ )  $t(812) = -4.612$ ,  $p < .000$ . Durations of /e/ ( $M = 132.65$ ;  $SD = 50.13$ ) and /a/ ( $M = 149.1$ ;  $SD = 52.83$ ) also differed significantly, with  $t(816) = -4.568$ ,  $p < .000$ . Durations between /i/ and /a/ were not tested, as a significant difference is assumed, given the previous results.

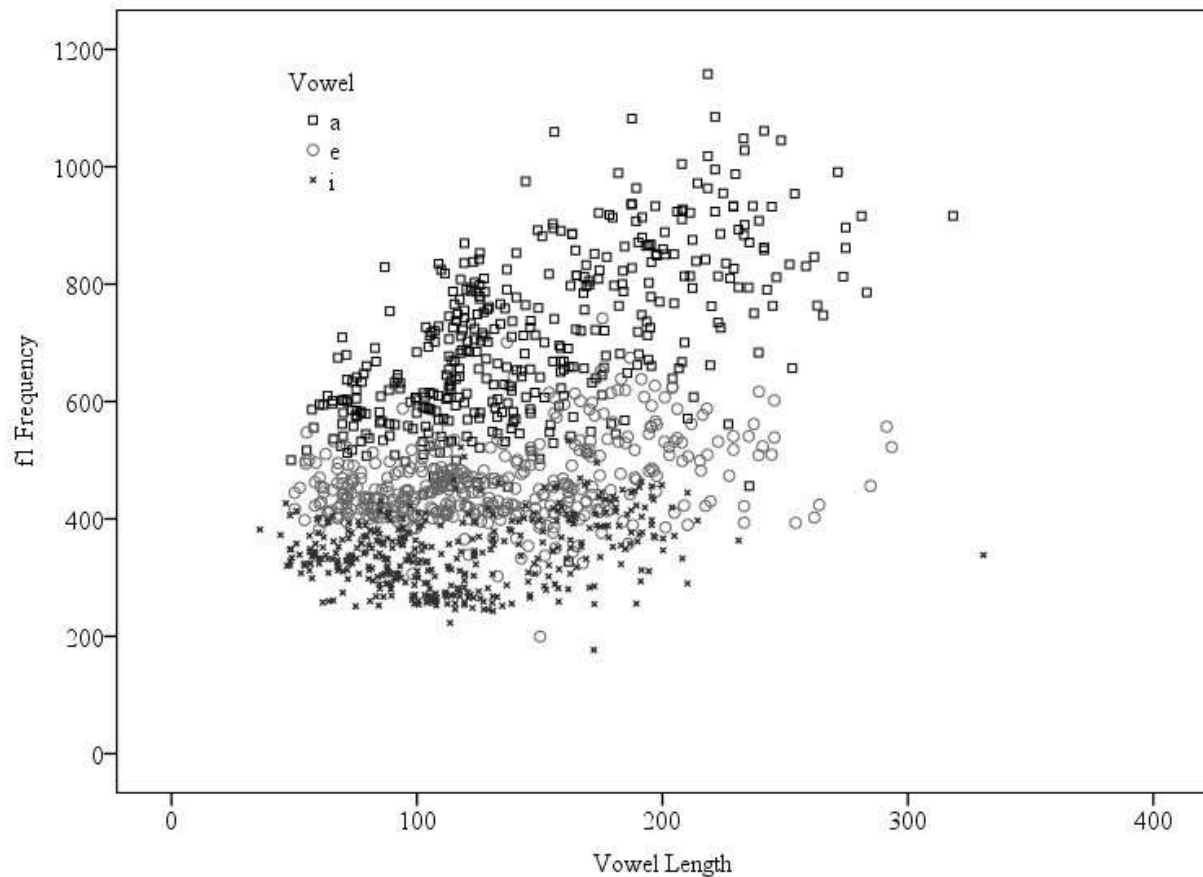
Significant f1 differences occurred between /i/ ( $M = 345.76$ ;  $SD = 64.46$ ) and /e/ ( $M = 462.36$ ;  $SD = 67.73$ )  $t(812) = -25.162$ ,  $p < .000$ , and between /e/ ( $M = 462.36$ ;  $SD = 67.73$ ) and /a/ ( $M = 710.72$ ;  $SD = 138.102$ )  $t(816) = -32.555$ ,  $p < .000$ . Like duration, f1 differences between /i/ and /a/ did not need to be tested. For complete results from these tests, see *Appendix B*.

### 2.4.3 Within-Categories

Using data points for both duration and F1, Pearson's correlations were performed speaker-by-speaker to find correlations within categories, which can be seen in full, in *Appendix C*. Overall, there was no correlation within-categories. However, tests of Pearson's R showed a

weak negative correlation for participants 1 and 3 for [a], and 3 again for [e], and a weak positive correlation for participant 4 for [a]. I cannot explain participant 4's single weak positive correlation, and conclude that it is likely due to an anomaly in the data. The overall lack of a correlation and potential weak negative correlation coincides with Toivonen et al. (2014). For a graphical representation of the between and within categories results, see Figure 4 below:

Figure 4: Within and between-category results for Study 1



Vowels differ significantly between categories but not within, suggesting Chilean Spanish is controlled. Vowel Length is in milliseconds, and f1 Frequency is in Hz.

## 2.5 Discussion

Between and within category analyses indicate that the mechanical explanation is insufficient for explaining DHC in Chilean Spanish, and that it is likely specified in the phonology. These findings are typical for this type of experiment (such as in Toivonen et al., 2014), and do not come as a surprise. Notwithstanding, these results are important evidence for DHC being a controlled phenomenon.

Solé and Ohala (2010) infer that DHC increases vowel distinctiveness, but as argued by Jakobson (1968), may be irrelevant in languages with highly contrastive vowels. According to Hualde (2005) and Payne (2008), Spanish has 5 distinct and highly contrastive vowels. However, as demonstrated, DHC in Chilean Spanish occurs similarly in English and Swedish (Toivonen et al., 2014), which have richer vowel inventories (Payne, 2008). Therefore, other factors must also be involved.

### 3. Study 2: Perception

Because of the findings from Lisker (1974), Tauberer and Evanini (2009), Solé and Ohala (2010) and Study 1 of this thesis, there is reason to believe that cross-linguistic DHC is not exclusively due to physiology. Questions arise as to why it is then cross-linguistic. The explanation may instead be found in the *perception* of duration. Perhaps, for instance, low vowels sound longer than high vowels, in addition to being measurably longer. To understand how DHC perception may play a role, it is important to understand how individuals use duration in speech.

Speakers may have preconceived notions of which vowels are longer than others. Bohn and Polka (2001) and Eilers et al. (1984) show that the ability to distinguish linguistic stimuli based on duration alone develops pre-linguistically, and is likely used throughout one's linguistic development. Because of this, speakers may be able to generalize these intuitions to unfamiliar forms (i.e. from unfamiliar languages).

In addition, there may be an incongruence between how speech sounds are perceived and how they measurably occur. Gussenhoven (2007) reports that high vowels are perceived as longer than low vowels. This finding is deemed to be due to participants' compensatory strategies for other articulatory side effects, such as changes in intensity in shifting contours.

Study 2 aims to investigate if English speakers have intuitions of vowel duration. Further exploratory measurements will be taken on participants' Response Times (RT), and Inter-Stimulus Interval conditions will be present to provide evidence for possible cognitive levels where duration processing may occur. Evidence will contribute to knowledge on the extent to which nuances in vowel duration are perceived, and the extent to which any generalizations can

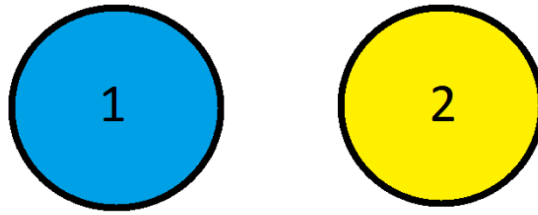
be made to unfamiliar sounds. Further details on previous research will be provided throughout this section.

### 3.1 Procedure

Upon signing relevant consent forms, participants performed a 2-Alternative Forced-Choice (2-AFC) task on a computer, where they directly interacted with the experimental interface by pressing keys corresponding to a choice. The experiment took on average 15 minutes to complete.

Participants were presented with instructions, which re-iterated the consent forms. After pressing a key to continue, they were presented with auditory stimuli consisting of a minimal pair (e.g. /bɪg/ - /bæg/), and were instructed to choose the longer sounding word. Each sound clip was represented visually with a number inside of a circle, as shown below in Figure 5. The first word in the pair was represented by ①, and the second with ②. Participants pressed the “z” key (covered in blue tape) if they thought the first word was longer, and the “m” key (yellow tape) if they thought the second was longer. The tape colour corresponded to the colour and orientation of the visual display in order to make the task as straightforward as possible. For each participant, the presentation volume was maintained at a comfortable listening level.

Figure 5: Main stimuli presented to participants



Due to the large number of stimuli, participants were allowed rest periods in between trials for as much time as required. This ensured they did not become bored during the task and choose items randomly to speed up the experiment, thus increasing the results' validity. To transition from one stimulus to the next, a fixation cross was flashed on a blank screen for 50 milliseconds as that is the minimum amount of time required for humans to notice a change has occurred (Meador et al., 2002). In doing this, participants noticed a change on the screen, and the experiment's length was minimized. Upon completion, participants were presented with a thank you message, and instructed to collect their \$5 gift card.

### 3.2 Participants

31 native English speakers of normal hearing completed the study. Disciplinary background should not affect the results, but participants were required to be a native speaker of English or have native proficiency, with no Swedish language background. Each participant was provided with a form detailing the instructions of the experiment and with a consent form (see Appendix I). Participants were compensated with a \$5 Tim Horton's gift certificate, regardless of whether or not they completed the experiment.

### 3.3 Measures

#### 3.3.1 Choice Task

Participant choices for minimal pairs were recorded, and results from these choices will be compared with each other to observe any potential trends in the data. For instance, comparisons could show any tendency for participants to choose words with low vowels as longer or shorter. If, for example, participants heard the word pair /bit – bæt/, they would be required to choose one of the words, and the amount of times that the particular word and words with similar vowel properties were chosen would be recorded. This type of data will be in the form of percentage of choices (e.g. low vowels were selected 50% of the time). Results from these measurements are important for indicating if participants have underlying intuitions for vowel length based on f1 height.

#### 3.3.2 Response Time

For each of the experimental stimuli, Response Times (RT) were recorded, and are of interest as they indicate participants' cognitive processing speeds when they are presented with stimuli and are required to make a choice (Jensen, 2006). These types of RT measures are known as *Choice Reaction Times* (CRT) (Sternberg, 1969). Knowing processing speed is important, as it is a meaningful index of how much attention is paid to the task, and if any interference is causing delayed responses (Stroop, 1935). RT is also a reliable measure of cognitive load (Brunken, Plass, & Leutner, 2003, p.53), and is thus important to understanding how much effort participants put towards making a choice.

### 3.3.3 Inter-Stimulus Intervals

Inter-Stimulus Intervals (ISI) are differences in the amount of time between two stimuli in minimal pairs. According to Werker and Logan (1985), playing auditory stimuli at different intervals (i.e. 500 milliseconds vs. 1000 milliseconds apart) activates different levels of cognitive processing (p.35). When given a discrimination task with a short ISI (500 milliseconds apart), participants displayed *acoustic*-level processing of non-native stimuli (Werker & Logan, 1985, p.42), meaning they encoded only auditory information into their short-term memory ( Craik & Lockhart, 1972, p.675). Information is briefly stored in a maintenance loop (McLeod, 2007). In a choice task, participants who have encoded their information acoustically would directly compare each stimulus to the other (Pisoni, 1975, p.8). This type of processing is associated with lower working-memory demands, meaning participants are sensitive to more detailed acoustic information (Werker & Logan, 1985).

Phonetic processing occurs when the ISI is longer, at approximately 1000 milliseconds, and involves accessing stored information in one's mental lexicon. When making choices on a phonetic level, Liberman et al. (1957; in Pisoni, 1975, p.1) reported that participants discerned whether or not two stimuli presented were phonetically the same or different based on their familiarity with speech sounds. In this level of processing, listeners would compare forms they encountered to the most similar matches in their lexicons.

In short, stimuli presented closely together in a short ISI (~500 ms) encodes as acoustic information, as opposed to a long ISI (~1000 ms), which encodes as phonetic information. Furthermore, a long ISI enables participants to compare two words by accessing previously stored sound information in their mental lexicons, while a short ISI causes participants to directly



compare words in a pair to one another.<sup>10</sup> In this study, a short ISI will include stimuli approximately 500 milliseconds apart, and a long ISI will be approximately 1000 milliseconds (one second) apart.

### 3.4 Materials

The stimuli for the experiment was recorded by myself (for the English word list) and by my supervisor Dr. Ida Toivonen (for the Swedish word list) using a Solid-State TASCAM recording device. Sound files were modified in the PRAAT acoustic software (Boersma & Weenink, 2014) using scripts for intensity (Vicenik, 2009b), which was controlled to 70 dB and pitch (Vicenik, 2009a), which was controlled to 262 Hz for the Swedish stimuli and 100 Hz for the English stimuli, as the Swedish stimuli were recorded by a female with a higher pitch than that of English stimuli recorded by a male. Stimuli duration information was gathered using a script by Mietta Lennes (2011), which also gathered f1 and f2 information, found in Table 1 below.

Table 1: Mean stimuli results for f1 height and f2 backness (Hz), and duration (milliseconds).

Mean	Swedish				English		
	/y/	/ø/	/œ/	/u/	/ɪ/	/ɛ/	/æ/
f1	368.52	611.34	403.93	531.42	416.29	572.31	697.38
f2	2070	1539.17	1655.29	1384.69	1916.12	1680	1584.59
Duration	276.63	293.82	261.23	268.12	144.81	162.33	213.87

<sup>10</sup> Though a third level of processing has been suggested in Werker and Logan (1985), the dual-processing concept used by Fujisaki and Kawashima (1969) will suffice in this experiment to remain straightforward in a study where ISI research is secondary to the main research.

The Swedish front and central vowels /y/ and /ø/ (e.g. in the Swedish “nytt” and “nött”), and central and back vowels /ə/ and /ɔ/ (as in “snudd” and “snodd”) were used, where both pairs include a high and mid vowel. There is disagreement in the literature concerning which phonetic characters should be used for Swedish vowels, with numerous Swedish dialects being an added issue. All vowels used were phonologically short. Each pair has a vowel with a high f1 and a low f1. Dr. Toivonen does not speak standard Swedish, but instead speaks a dialect found in the Åland islands of Finland. See Leinonen (2010) for more information of vowel quality in Swedish spoken in Finland.

Pair lists were arranged in Microsoft Excel and exported into the PsychoPy experimental presentation software (Pierce, 2007). ISIs were directly specified in PsychoPy, which also randomized the order of presentation. The experiment itself was carried out on an Acer Aspire E1-510P-2822 laptop.

### 3.5 Design

Study 2 is a 2-AxB Alternative Forced Choice Task, where participants choose which word in a minimal pair is the longest. There is no “same duration” option, meaning participants are “forced” to choose one word over the other. This task is an adaptation of a similar study performed by Lehnert-LeHouillier (2010), where participants of varying linguistic backgrounds listened to words containing vowels of different length.

Participants will encounter both English and Swedish stimuli, and will be presented both with meaningful words, and nonsense (nonce) words (see Appendices *D* and *E* for the complete word list). The list of nonce words accidentally included some words that were, in fact real words. A future study will exclude these words to make sure that the (lack of) differences between real and nonce words, are, indeed, as discussed in this thesis. Words in minimal pairs

(such as /bɪg/ - /bæɡ/) will retain their original durations, and in another condition be altered to have equal durations. Equal durations will be the average duration of both words in the minimal pair, and are thus pair-specific. More information on vowel manipulations can be found in the PRAAT user manual (Styler, 2013) regarding duration (p.34), intensity (p.35), and pitch (p.32). Minimal pairs from the original duration condition will be reused in the modified duration condition, as in Tables 2 and 3 below. All minimal pairs were repeated in reversed order (first word with a lower vowel, second with a higher vowel), to control for ordering effects. All conditions (English/ Swedish; modified/ unmodified duration; original/ reversed order) will be presented with a short and long ISI in alternating trial blocks. Participants were not informed of the modified duration condition and the ISI differences.

Table 2: Experimental Conditions for the English stimuli. Numbers in each slot represents the number of minimal pairs for each condition.

	No Modification	Duration Modified
Real Words	9	9
Nonce Words	12	12

Table 3: Experimental Conditions for the Swedish stimuli. Numbers in each slot represents the number of minimal pairs for each condition.

	No Modification	Duration Modified
Real Words	12	12
Nonce Words	12	12

Though the experiment was divided into five trial blocks, they were not separated by condition apart from ISI. A total of 180 stimulus pairs were used, which consisted of 90 minimal pairs that were copied and had their order reversed to control for ordering, effectively doubling the list. Stimuli from each condition was presented in randomized order in each trial. Each block consisted of 40 pairs randomly selected from the main stimulus list, which had been randomized further in Microsoft Excel.

### 3.6 Predictions

Though high vowels are measurably shorter than low vowels, there is still uncertainty to whether or not they also *sound* shorter. To my knowledge, no previous studies have focused on the outcome of a forced-choice task similar to this study. However, experiments with concepts similar to what is expressed in this research can lead me to consider a given set of patterns that the data are likely to follow. Different scenarios pertaining to different elements of the experiment will be detailed in the sections below.

#### 3.6.1 Vowel Pairs

Participants are expected to perceive low vowels as sounding longer than high vowels. As low vowels are measurably longer than high vowels, listeners may directly perceive a similar correlation. Though little research has investigated how vowel height affects the perception of duration in words, listeners have been shown elsewhere in the literature to perceive changes in vowel length in languages that do not employ contrastive duration. In Raphael (1972), participants successfully discriminated between unreleased English voiced and unvoiced stops by

judging the duration of the preceding vowel (p.1301). Just as vowels preceding voiced stops tend to be measurably longer than before unvoiced stops (Umeda, 1975), participants in Raphael (1972) perceived them as longer. Perhaps then, duration perception could occur similarly with vowel height. However, there is some debate over this in the literature. In this experiment, the English stimuli consists of pair words that only contain voiced codas, to ensure that effects from coda voicing are not an issue.

Gussenhoven (2007) performs a study where listeners are asked to judge on a 7-point scale how long a vowel is compared to an anchor vowel. In his study, participants judged high vowels as longer than low vowels (p.13). Further evidence for this can be observed in vowel chain shifts, whereby long vowels in a language will tend to raise and become higher, and short vowels will become lower (Labov, 1994, in Jacewicz, Fox, & Salmons, 2006, p. 287). However, these studies address different questions than this thesis, as they discuss how vowel length affects  $f_1$ , and not how changes in  $f_1$  affect perceived length.

Participants are expected to choose low vowels as longer to a greater extent in English than in Swedish, as they have no previous exposure to Swedish vowels. In Flege, Bohn, and Jang (1997), listeners hearing an unfamiliar language perceived its speech sounds with constraints from their own native language (p.2548). It is important to keep in mind that listeners can typically distinguish vowel durations that are 10 to 40 milliseconds apart (Lehiste, 1970; in Catford, 1977, p.196).

The Swedish stimuli is of great interest as the words contain vowels /y/, /ø/, /œ/, and /ʊ/; unknown to Canadian English speakers unfamiliar with Scandinavian languages. As participants have been controlled to have no Swedish linguistic background, they are all assumed to be performing choice tasks with no previous knowledge of the vowels.

### 3.6.2 Natural vs. Modified Duration

As DHC occurs cross-linguistically (Lehiste, 1970) and the ability to distinguish linguistic duration develops early in human development (Eilers et al., 1984; Bohn & Polka, 2001), participants are expected to categorize the longer word in a pair based on the relevant vowel's duration, when the word pairs have their *natural* durations intact. When the durations are modified, their abilities to categorize vowels may remain intact if they have previous exposure to the vowels. In the case of English, they do, and in the case of Swedish, they do not.

As this is a forced-choice task, participants are required to categorize vowels to the best of their ability. Diminished performance could thus be measured by an increase in the time it takes participants to complete the choice (RT) and/or a decrease in participants' accuracy. Performance differences between normal and modified duration word pairs for Swedish stimuli are expected to follow a similar pattern to what is found in the English condition, albeit with a potentially longer RT, or less of a pattern in participants' choices.

### 3.6.3 Real vs. Nonsense Words

Performance across both durational conditions regarding English nonce words is expected to be similar to that with meaningful words, as both nonce and meaningful words follow the same phonological rules and constraints that dictate their composition. For instance, Wang, Lehiste, Chuang, and Darnovsky (1976) found that participants perceived vowel stimuli in a consistent pattern, regardless of whether or not the forms they were presented with contained meaning. This being said, any performance differences that were to occur would likely be due to participants' unfamiliarity with the stimuli.

Though nonce words are used for the Swedish condition, no performance difference is expected between real and nonce Swedish words, as the speaker is presumed to be unfamiliar with Swedish. Modified and non-modified nonce words are still included in order to balance the English word lists, and to control for any potential interference to the production of the nonce stimuli.

#### 3.6.4 Short vs. Long ISI

Though different cognitive processing levels are activated by ISI length, both acoustic and phonetic levels of processing should yield similar vowel length categorizations. In phonetic encoding (long ISI), participants would compare each word to their linguistic knowledge, whereas in acoustic processing (short ISI), each word's duration would be compared to the other in the pair. In this way, different ISIs should alter the underlying way in how vowels are categorized, but not whether or not the vowels can be categorized at all.

### 3.7 Results

The below subsections will detail the main results of study 2. For a complete record of the data, see Appendix F. All tests of significance will be explored with  $\alpha$ -value of .05. If an  $\alpha$ -value of .01 is instead applied, most effects remain significant, except for the real-nonce data and natural vs. modified Swedish vowels between pair orders.

#### 3.7.1 Subjective Reports

Participants enjoyed the experiment, and found the task simple and straightforward. As each trial block contained a substantial number of stimuli, participants welcomed the breaks

given to them. Though some participants reported the experiment as repetitive, that aspect is neither avoidable nor consequential to the validity of the data.

Participants reported low vowels as sounding longer than high vowels. However, they also reported suspicions that many pairs were controlled to be the same length. As the task was forced-choice by nature, they proceeded to choose the “correct” word to the best of their ability. The reports did not differ with regards to whether the words were real or nonce; English or Swedish. Some participants reported that upon hearing word pairs they had already encountered, they relied on their previous choice to make a choice, instead of focusing entirely on the word pair at hand. This was partially controlled for by presenting repetitions with words in switched-order. The fact that participants were still able to rely on previous choices means that this effect was not entirely controlled, and future research should focus on removing this effect.

Some participants reported being unsure of which word was longer, and proceeded to guess. Whether or not these guesses are intuitive in nature is for the data to show, but means that for some participants, word choices were unconscious. Participants were as a whole unable to guess that the second language was Swedish, confirming that no Swedish speakers were tested. Observations consisted entirely of the decision-making component of the experiment, and participants were unaware that they were being measured for RT, and that conditions were separated by ISI.

In the following sections, each subsection in the Results corresponds its respective subsection in the Discussion.



### 3.7.2 Choice Data

Due to differences in the total number of choices for each vowel pair, tests of chi-squared ( $\chi^2$ ) cannot completely account for significance. As such, the  $\chi^2$  effect size  $\phi$  will be included to indicate how meaningful the findings are.

Paired-samples t-tests found no significant differences between English word pairs in their High-Low Vowel order (i.e. big - bag) ( $M = 80.99$ ,  $SD = 12.33$ ), and reversed (low vowel-high vowel) word order (i.e. bag - big) ( $M = 77.79$ ,  $SD = 14.51$ ),  $t(18) = 0.9144$ ;  $p = 0.3726$ ,  $d = .75$ . Participants' choices have thus been collapsed across word orders, meaning that length choice data in the subsections below will be the averages of word pairs in both orders. For results not collapsed across word order, see Appendix F.

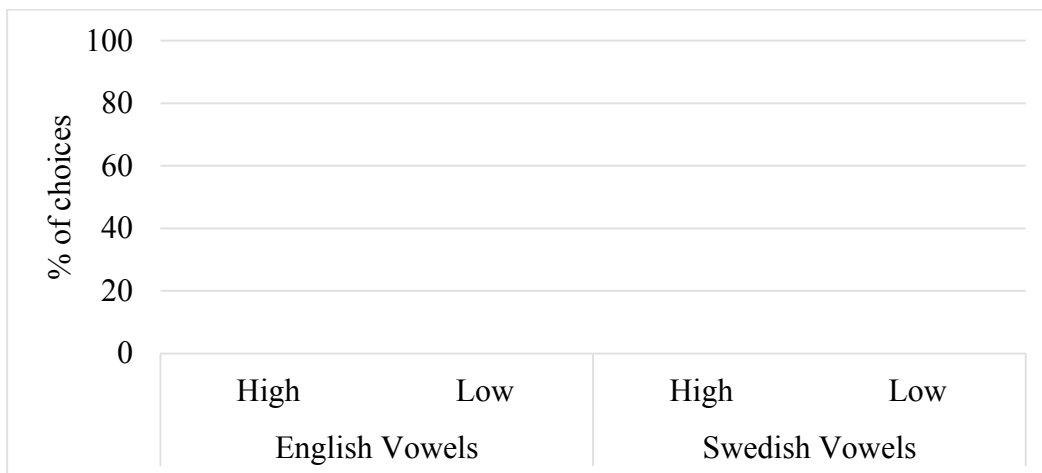
For Swedish however, participants categorized low vowels as long 15% more when pairs were in their normal order (first word: high vowel, second word: low vowel) ( $M = 75.84$ ,  $SD = 5.09$ ) than when they were in reversed order ( $M = 60.85$ ,  $SD = 6.99$ ),  $t(6) = 3.9627$ ;  $p = .0074$ ,  $d = 3.24$ . In addition, normally-ordered Swedish pairs had an  $\chi^2$  effect size of  $\phi = .51$ , whereas reverse-order pairs had a mean effect size of  $\phi = .23$ . As the normally-ordered pairs' effect sizes are on average more than twice the size than that of reverse-order pairs, a stronger effect is observed for word pairs with a high vowel followed by a low vowel.

Pair ordering in Swedish thus has a significant effect on duration perception. Participants unfamiliar with Swedish rely on additional cues (such as ordering) to make judgements on vowel quantity. Pinker (1994, p.167) and Cooper and Klouda (1995, p.335) noted a tendency for English speakers to prefer reduplicated (minimal) pairs where the vowel of the first word is high, and the vowel of the second is low (i.e. "zig-zag). This will be explored further in the Discussion section.

Across all conditions, participants categorized words with English low vowels (i.e. /ɛ/, /æ/) ( $n = 1995$ ) as longer 79.83% more than high vowels ( $n = 528$ ),  $\chi^2(1, N = 2523) = 852.99, p < .0001, \phi = 0.58$ . Similarly, Swedish mid vowels (i.e. /ʊ/, /ø/) in the normal order ( $n = 958$ ) were chosen 75.88% more than corresponding high vowels (i.e. /ə/, /y/) ( $n = 313$ ),  $\chi^2(1, N = 1271) = 327.32, p < .0001, \phi = 0.51$ . In the reversed order however, lower vowels ( $n = 803$ ) were categorized as longer 61.38% more than high vowels ( $n = 499$ ),  $\chi^2(1, N = 1302) = 70.98, p < .0001, \phi = 0.23$ , indicating a difference between normal and reverse pairs of 14.5%.

English low vowels were selected as longer 10.09% more than Swedish low vowels across both orders, perhaps due to participants' unfamiliarity with the Swedish vowel system. This pattern of results is shown below, in Figure 6.

Figure 6: English vs. Swedish Length Choices



### 3.7.2.1 Vowel Pairs

For English vowel pairs, a paired-samples t-test found participants selected lower vowels as longer significantly less in /ɪ - ε/ pairs ( $M = 64.49$ ,  $SD = 9.19$ ) than in /ɪ - æ/ pairs ( $M = 87.03$ ,  $SD = 5.27$ ),  $t(15) = 10.0123$ ,  $p < .0001$ ,  $d = 5.17$ , or /ε - æ/ pairs ( $M = 88.39$ ,  $SD = 4.22$ ),  $t(15) = 8.9791$ ,  $p < .0001$ ,  $d = 4.64$ . However, /ɪ - æ/ pairs did not differ significantly from /ε - æ/ pairs  $t(15) = 0.878$ ,  $p = .3938$ ,  $d = .45$ .

Significance was found between words within the same pair. A summary of within word pair comparisons is shown below in Table 4. These mean values follow similar patterns in all of the different English conditions.

Table 4: Summary of within-pair comparisons for English vowels

Vowel Pair	# High Vowel	# Low Vowel	% Chosen	Chi-Squared	Effect Size ( $\phi$ )
/ɪ - ε/	338	642	64.49	94.3	0.31
/ɪ - æ/	112	751	87.03	471.14	0.74
/ε - æ/	78	602	88.39	403.79	0.77

All findings are significant,  $p < .0001$ . Low vowels in the pair are shown in bold.

For Swedish normally-oriented pairs, a paired-samples t-test found participants select low vowels as longer in /y - ø/ pairs ( $M = 75.62$ ,  $SD = 3.04$ ) approximately as much as in /ø - u/ pairs ( $M = 76.31$ ,  $SD = 5.44$ ),  $t(4) = 0.4222$ ,  $p = 0.6946$ ,  $d = .42$ . For reverse-ordered pairs, low vowels in /ø - y/ pairs ( $M = 61.93$ ,  $SD = 7.91$ ) were selected approximately the same amount as in /u - ø/ pairs ( $M = 60.85$ ,  $SD = 3.76$ ),  $t(4) = 0.2216$ ,  $p = 0.8355$ ,  $d = .22$ . Overall, low vowels

in normally-ordered pairs were chosen significantly 14.38% more of the time than in reverse-ordered pairs  $t(9) = 5.6079, p = 0.0003, d = 3.74$ . A summary within word pair comparisons is shown below in Table 5.

Table 5: Summary of within-pair comparisons for Swedish vowels

Vowel Pair	# High Vowel	# Low Vowel	% Chosen	Chi-Squared	Effect Size ( $\phi$ )
/y - <b>ø</b> /	209	628	75.62	209.75	.5
/ <b>ø</b> - <b>ɔ</b> /	104	330	76.31	117.69	.52
/ <b>ø</b> - y/	266	447	61.93	45.95	.25
/ <b>ɔ</b> - <b>ø</b> /	233	356	60.85	25.69	.21

All findings are significant to  $p < .0001$ . Low vowels in the pair are shown in bold.

### 3.7.2.2 Natural vs. Modified Duration

Participants' choices did not differ in either language with regards to whether or not the vowels kept their natural durations ( $M = 80.88, SD = 12.55$ ), or were modified to have identical durations ( $M = 78.79, SD = 11.47$ ),  $t(11) = 1.1552, p = .2725, d = .7$ . Participants chose English low vowels with natural durations intact ( $n = 947$ ) 80.88% more often than corresponding high vowels ( $n = 223$ )  $\chi^2(1, N = 1170) = 448.01, p < .0001, \phi = 0.62$ . Similarly, modified low vowels ( $n = 1048$ ) were chosen 78.79% more often than high vowels ( $n = 305$ )  $\chi^2(1, N = 1353) = 408.01, p < .0001, \phi = 0.55$ . Comparisons of natural and modified vowel duration results are shown below in Table 6.

Table 6: English length choices between natural and modified duration conditions

	Natural Duration			Modified Duration		
	/ɪ - ɛ/	/ɪ - æ/	/ɛ - æ/	/ɪ - ɛ/	/ɪ - æ/	/ɛ - æ/
% Chosen	64.68	88.85	89.12	65.12	84.13	87.09
#V1 Chosen	137	50	36	201	62	42
#V2 Chosen	264	380	303	378	371	299
Effect Size $\varphi$	.32	.77	.79	.31	.71	.75

The bold vowel in each pair refers to the measurably lower one. “% Chosen” refers to how much the lower vowel was chosen. “#V1 Chosen” refers to the amount of times the first (higher) vowel in the above pairs was selected, and “#V2 Chosen” refers to the second (lower) vowel. All effects are significant,  $p < .0001$ .

In Normally-Ordered pairs, Swedish low vowels with natural durations ( $M = 74.32$ ,  $SD = 0.95$ ) were chosen significantly (3.55%) less than modified low vowels ( $M = 77.87$ ,  $SD = 0.01$ ),  $t(2) = 6.5378$ ,  $p = 0.0226$ ,  $d = 9.25$ . In Reverse-Order pairs Swedish low vowels with natural durations ( $M = 63.36$ ,  $SD = 5.29$ ) were chosen a non-significant 3.95% more than modified low vowels ( $M = 59.41$ ,  $SD = 4.22$ ),  $t(2) = 0.7188$ ,  $p = .5469$ ,  $d = .77$ .

In the Normal Order, participants chose low vowels with natural durations intact ( $n = 476$ ) as longer 74.32% more than high vowels ( $n = 175$ ),  $\chi^2(1, N = 651) = 139.17$ ,  $p < .0001$ ,  $\varphi = .46$ , and chose corresponding modified low vowels ( $n = 482$ ) 77.87% more than modified high vowels ( $n = 138$ ),  $\chi^2(1, N = 620) = 190.87$ ,  $p < .0001$ ,  $\varphi = .56$ .

In the Reverse Order, participants chose low vowels with natural durations intact ( $n = 411$ ) as longer 63.36% more than high vowels ( $n = 240$ ),  $\chi^2(1, N = 651) = 44.92$ ,  $p < .0001$ ,  $\varphi = .26$ , and chose corresponding modified low vowels ( $n = 392$ ) 59.41% more than modified high

vowels ( $n = 259$ ),  $\chi^2(1, N = 651) = 27.17, p < .0001, \phi = .2$ . Comparisons of natural and modified vowel duration results for Swedish are shown below in Table 7.

Table 7: Swedish length choices between natural and modified duration conditions

	Natural Duration		Modified Duration	
	/y - ø/	/ø - u/	/y - ø/	/ø - u/
% Chosen	<b>73.36</b> , 68.66	<b>75.27</b> , 58.46	<b>77.86</b> , 55.19	<b>77.88</b> , 63.63
#V1 Chosen	<b>119</b> , 108	<b>56</b> , 132	<b>90</b> , 158	<b>48</b> , 101
#V2 Chosen	<b>315</b> , 233	<b>161</b> , 178	<b>313</b> , 214	<b>169</b> , 178
Effect Size $\phi$	<b>.45</b> , .37	<b>.48</b> , .15	<b>.55</b> , .15	<b>.56</b> , .28

Each result is composed of normally-ordered pair results on left (in bold), followed by reverse-ordered pairs on right, separated by a comma

### 3.7.2.3 Real vs. Nonsense words

Participants were selected to not have any background in Swedish, so they are presumed to not understand the Swedish words they heard. Because of this, only English has the real/nonsense (nonce) contrast.<sup>11</sup>

In the real word condition, words containing low vowels ( $n = 645$ ) were chosen 77.1% of the time in contrast to high vowels ( $n = 187$ ),  $\chi^2(1, N = 832) = 252.12, p < .0001, \phi = .55$ . In the nonce condition, low vowels ( $n = 1364$ ) were chosen 80.66% more often than high vowels ( $n = 341$ ),  $\chi^2(1, N = 1705) = 613.8, p < .0001, \phi = 0.6$ ;  $t(19) = 2.8843, p = 0.0242, d = 1.12$ . Table 8

<sup>11</sup> A paired-samples t-test for the Swedish was performed to ensure there was no effect, and indeed no significance was found in vowel length choices,  $t(12) = 0.8516, p = .4111, d = 0.49$  or RT,  $t(935) = 1.2774, p = .2018, d = .01$ .

below summarizes these findings on a vowel-by vowel basis. This goes against my hypotheses as participants were predicted to have greater trouble making choices for nonce words than real words. Because of this, the fact that the proportion of low vowel choices in the nonce condition is not significantly less than the real word condition is of great interest, and will be elaborated on in the Discussion. It should be noted that though these differences are significant, they are still quite small, and may not necessarily be *meaningful*.

Table 8: English real vs. nonce vowel choices

	Real Words			Nonce Words		
	/ɪ - ɛ/	/ɪ - æ/	/ɛ - æ/	/ɪ - ɛ/	/ɪ - æ/	/ɛ - æ/
% Chosen	64.13	85.22	85.28	65.67	87.76	90.92
#V1 Chosen	115	36	36	223	76	42
#V2 Chosen	214	207	224	428	544	392
Effect Size $\phi$	.30	.70	.72	.31	.75	.81

### 3.7.2.4 Short vs. Long ISI

In English, low vowels in the short ISI condition ( $M = 78.81$ ,  $SD = 12.14$ ) were not chosen significantly more than in the long ISI condition ( $M = 79.13$ ,  $SD = 13.45$ ),  $t(18) = 0.1419$ ,  $p = 0.8888$ ,  $d = .07$ .

When English word pairs were presented with a short Inter-Stimulus Interval (ISI) (500 ms), low vowels ( $n = 1076$ ) were chosen 78.81% more than high vowels ( $n = 301$ ),  $\chi^2(1, N = 1377) = 436.18$ ,  $p < .0001$ ,  $\phi = .56$ . With a long ISI (1000 ms), low vowels ( $n = 919$ ) were

chosen 79.13% more than high vowels ( $n = 227$ ),  $\chi^2(1, N = 1146) = 417.86, p < .0001, \varphi = .60$ .

Table 9 below summarizes these findings on a vowel-by vowel basis.

Table 9: English Choices for Long and Short ISI

	Short ISI (500 ms)			Long ISI (1000 ms)		
	/ɪ - ɛ/	/ɪ - æ/	/ɛ - æ/	/ɪ - ɛ/	/ɪ - æ/	/ɛ - æ/
% Chosen	62.02	85.1	87.3	67.78	87.88	89.91
#V1 Chosen	190	62	49	148	50	29
#V2 Chosen	325	399	352	317	352	250
Effect Size $\varphi$	.26	.73	.76	.36	.75	.79

In High-Low-vowel Swedish pairs, whether the ISI was short ( $M = 74.659, SD = 0.1$ ) or long ( $M = 61.54, SD = 0.23$ ) had no significant effect on participants' choices  $t(1) = 7.782, p = 0.0813, d = .16$ . Low vowels ( $n = 319$ ) were chosen 74.66% more than high vowels ( $n = 115$ ),  $\chi^2(1, N = 434) = 95.89, p < .0001, \varphi = .47$ . With a long ISI, low vowels ( $n = 555$ ) were chosen 78.88% more than high vowels ( $n = 158$ ),  $\chi^2(1, N = 713) = 221.05, p < .0001, \varphi = .56$ .

In Reverse-Ordered (Low-High vowel) Swedish pairs, low vowels in pairs with a short ISI ( $M = 62.2, SD = .79$ ) were chosen significantly more than vowels in pairs with a long ISI ( $M = 60.12, SD = 1.25$ ),  $t(3) = 2.0936, p = .0134, d = 2.42$ . Low vowels with a short ISI ( $n = 465$ ) were chosen 62.2% more than high vowels ( $n = 279$ ),  $\chi^2(1, N = 744) = 46.5, p < .0001, \varphi = .25$ . With a long ISI, low vowels ( $n = 338$ ) were chosen 60.12% more than high vowels ( $n = 220$ ),  $\chi^2$



(1,  $N = 558$ ) = 24.95,  $p < .0001$ ,  $\phi = .21$ . Table 10 below summarizes these findings on a vowel-by vowel basis.

Table 10: Swedish Choices for Long and Short ISI

	Short ISI (500 ms)		Long ISI (1000 ms)	
	/y - ø/ <sup>a</sup>	/ø - u/	/y - ø/	/ø - u/
% Chosen	61.39	<b>74.66</b> , 63.23	<b>77.42</b> , 61.4	<b>80.34</b> , 58.47
#V1 Chosen	89	<b>115</b> , 102	<b>94</b> , 89	<b>64</b> , 131
#V2 Chosen	159	<b>319</b> , 177	<b>309</b> , 159	<b>246</b> , 179
Effect Size $\phi$	.28	<b>.47</b> , .27	<b>.53</b> , .28	<b>.59</b> , .15

<sup>a</sup>Only reverse-ordered data is available in this condition

These results are interesting, as the opposite effect would be expected, given Werker and Logan's (1985) results, where stimulus pairs with a short ISI were shown to have low vowels that were chosen more often than with a long ISI. Here however, for English, short and long ISI results do not significantly differ from one another, differing by only 3.38%,  $t(22) = .6944$ ,  $p = 0.4947$ . In Swedish however, there appears to be a higher choice polarity when words are presented with a short ISI. This pattern of results will be further discussed in the Discussion.

### 3.7.3 Response Time Data

Response times were in general quite long for a length-categorization task such as this (see Weker & Logan, 1985), which could be the result of the nature of the task itself (e.g. hearing

one word presented after another), a variable in stimulus presentation (e.g. in PsychoPy), or other reasons not accounted for here.

No significant difference was between word pairs in their High V-Low V order (i.e. /big/-/bag/), and reversed word order (i.e. /bag/-/big/) for both English  $t(30) = -0.02706, p = 0.579$ , and Swedish  $t(30) = 0.023, p = 0.468$ . Therefore, both languages' data have been collapsed across pair order. As each stimulus was predicted in the experiment to take approximately 500 milliseconds to present, each response time value had 500 subtracted from it.

Overall, participants responded 250 milliseconds faster to English stimuli ( $M = 1137$  ms,  $SD = 12.43$ ) than to Swedish stimuli ( $M = 1388$  ms,  $SD = 1.55$ ). As English compares 3 vowels and Swedish 2, an unpaired t-test was employed, and found that response times to English stimuli are significantly shorter than to Swedish stimuli  $t(56) = 5.4541, p < .0001, d = 1.46$ . Thus English and Swedish will be analyzed separately using repeated-measures ANOVAs.

### 3.7.3.1 English RT Results

In the English data, Mauchly's test indicated that the assumption of sphericity had not been violated, thus no correction to the degrees of freedom were necessary. Significant results are shown below, in Table 11.

Table 11: English RT Results<sup>a</sup> – Repeated Measures ANOVA

Condition	F-Value	P-Value ( $\alpha = .05$ )	Partial Eta-Squared
Modified/ Natural Duration	4.226	.049	.123
Nonce/ Real	8.787	.006	.227
ISI	14.082	.001	.319
Vowel Pair	6.221	<.001	.398
Nonce/ Real x ISI	7.351	.011	.197
Mod./ Nat. Duration x V. Pair	3.645	.032	.108

<sup>a</sup> Only significant results are reported

Participants took 72 ms longer to make a choice with modified words ( $M = 1668$  ms,  $SE = 65$ ), than with pairs with their natural durations intact ( $M = 1596$  ms,  $SE = 66$ ).

Vowel categorization with nonce words ( $M = 1581$  ms,  $SE = 62$ ) was 103 ms faster than in real words ( $M = 1684$  ms,  $SE = 69$ ). Interestingly, this coincides with findings from the choice data, meaning participants were both more pointed in their choices in the nonce condition, and spent less time making these choices.

When pairs were presented with a short ISI ( $M = 1725$  ms,  $SE = 68$ ), participants spent 186 ms more on making a choice than when pairs had a long ISI ( $M = 1539$  ms,  $SE = 68$ ).

Participants chose low vowels in / $\varepsilon - \text{æ}$ / pairs ( $M = 1507$  ms,  $SE = 59$ ) most quickly, while taking 72 ms longer with / $\text{ɪ} - \text{æ}$ / pairs ( $M = 1579$  ms,  $SE = 73$ ). The most time was needed to make a decision in / $\text{ɪ} - \text{ɛ}$ / pairs ( $M = 1810$  ms,  $SE = 75$ ).

In sum, each experimental condition significantly affects response time. This differs from the English Length Choice data, where only whether the words were real or nonce affected performance. Though participants' overall ability to judge vowel length did not change, RT

differences across conditions indicate that there are different processing loads associated with the different conditions.

Figure 7 shows an interaction between semanticity and ISI, whereby real ( $M = 1823$ ,  $SE = 8.7$ ) and nonce ( $M = 1626$ ,  $SE = 6$ ) word RTs in the short ISI condition both decrease to approximately the same time in the Long ISI ( $M = 1539$ ,  $SE = 6.95$ ). Explanations for why this could occur will be further explored in the Discussion.<sup>12</sup>

Figure 7: RT for nonce/ real words across ISI

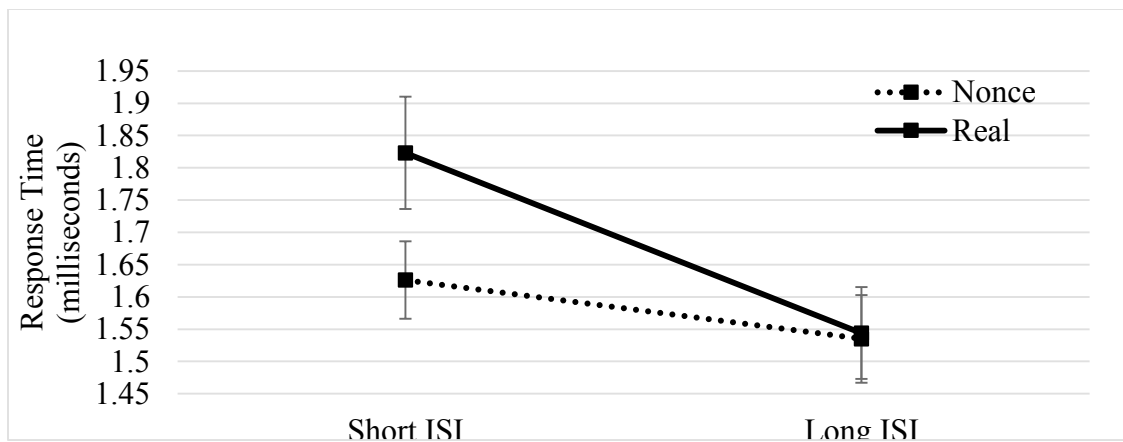
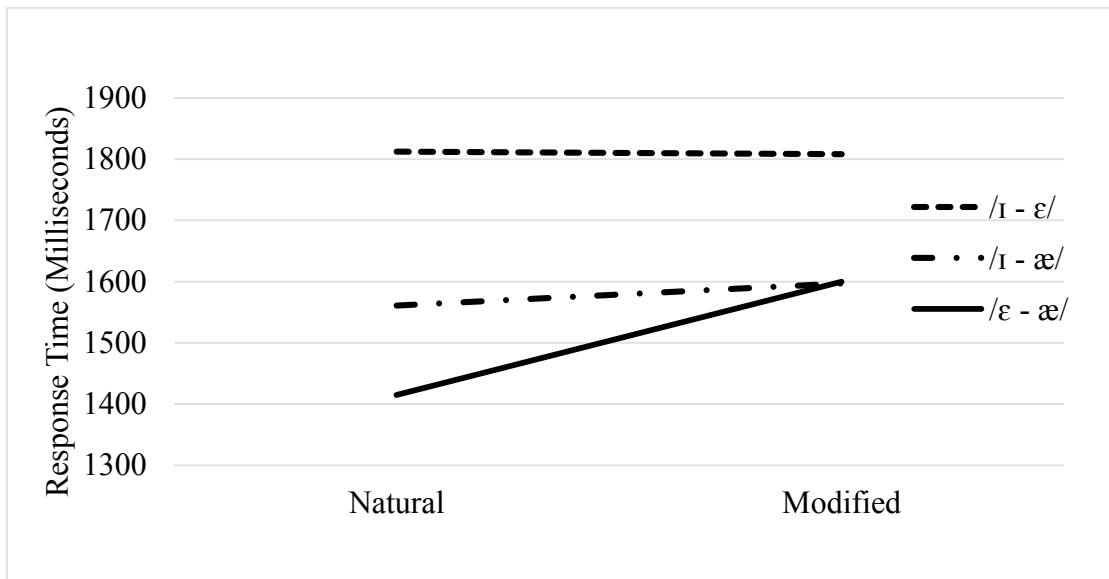


Figure 8 shows the interaction between which vowels were present in a pair, and whether or not those vowels' durations had been modified. No significant change occurred between /ɪ - ε/ pairs that kept their natural durations ( $M = 1812$  ms,  $SE = 8.9$ ) and those with modified durations ( $M = 1808$  ms,  $SD = 6.7$ ). A slight change occurs in /ɪ - æ/ pairs, whereby pairs with

<sup>12</sup> Because of an RT effect, ISI on semanticity was then tested for the Choice data. However, nonce words with a short ISI ( $M = 76.83$ ,  $SD = 14.25$ ) did not significantly differ from those with a long ISI ( $M = 78.54$ ,  $SD = 13.89$ ),  $t(10) = 0.2774$ ;  $p = .7872$ ,  $d = .18$ . Similarly, real words with a short ISI ( $M = 81.63$ ,  $SD = 11.87$ ) did not significantly differ from those with a long ISI ( $M = 81.81$ ,  $SD = 11.81$ ),  $t(11) = .0326$ ;  $p = 0.9746$ ,  $d = .59$ .

natural durations ( $M = 1561$  ms,  $SE = 6.9$ ) had a slightly shorter RT than pairs with modified durations ( $M = 1597$  ms,  $SE = 8.7$ ). The most meaningful change in RT occurs in / $\epsilon - \text{æ}$ / pairs, where natural duration pairs ( $M = 1415$  ms,  $SE = 6.5$ ) were responded to 185 ms more quickly than pairs from the modified condition ( $M = 1600$ ,  $SE = 6.6$ ).

Figure 8: Vowel pair RT across duration conditions



Though there is no clear indication of why this interaction would occur, speculatively, pair distinctiveness may have an effect on the ease at which participants are able to confidently judge vowel durations. As the shortest RT is observed for the most readily distinguished vowel pair in the English data, participants have little difficulty finding the longer vowel when the pair retains their natural durations. When the durations are controlled to be the same length, their performance does not decrease *per se*, but more time is required to make a choice.

### 3.7.3.1 Swedish RT data

A repeated-measures ANOVA found no significant effects anywhere in the Swedish RT data. This itself is interesting, as there are likely differences between how participants perceive familiar English vowels and unfamiliar Swedish vowels. A summary of the results for the different conditions is shown below, in Table 12.

Table 12: Swedish RT Results – Repeated Measures ANOVA

Condition	F-Value	P-Value ( $\alpha = .05$ )	Partial Eta-Squared
Modified/ Natural Duration	.031	.862	.001
ISI	.112	.740	.004
Vowel Pair	2.394	.097	.089

It should be noted in the above table that the Vowel Pair condition has significance values closest to approaching an  $\alpha$ -value of .05. Though not statistically significant, this does indicate that out of all the conditions, the backness of vowel pairs had the strongest effect on response times.

## 3.8 Discussion

The below subsections will discuss findings from each experimental measure (e.g. RT), and then integrate all findings to link them into a cohesive explanation of how speakers were perceiving vowel length in word pairs. Each subsection in the Discussion corresponds to its respective subsection in the Results.

### 3.8.1 Overall Trends

In both English and Swedish across conditions, participants selected low vowels as longer significantly more often than high vowels, indicating that contrary to Gussenhoven (2007) and Wang et al. (1976), low vowels are indeed both measurably longer, and sound longer. However, the stimuli used in this experiment were different to that in Gussenhoven (2007), where English words with unvoiced codas and Dutch stimuli were used. Dutch employs contrastive duration (Van Dommelen, 1982), which may affect perception, though (as indicated in the Introduction of this thesis) not production. In addition, Gussenhoven uses a different method. In his study, participants are asked to compare stimuli to an anchor vowel and judge the relative duration on a seven-point scale. Gussenhoven and Zhou (2013), using a similar method to that of Gussenhoven (2007), found that participants judged low vowels as longer than high vowel, similar to what was found in this thesis. Furthermore, Wang et al. (1976) used vowels in isolation, and not embedded in carrier words similar to stimuli found in this thesis.

Though this effect was significant for both English and Swedish, low vowels were selected as longer to a greater extent in English than in Swedish. Response-time evidence shows that participants made their English choices faster than Swedish choices, suggesting that they could more easily distinguish between English high and low vowels. Minagawa-Kawai, Mori, and Sato (2005) indicate with Near-Infrared Spectroscopy (NIRS) that native speakers of a language will have neural representations for that language's phonemes in cortical regions associated with speech processing (e.g. Wernicke's area) that speakers less exposed to a particular language do not (p. 1379). Though performance differences between languages are indeed expected, differences between Swedish high-low vowel pairs and English was miniscule, while the difference with Swedish reverse-ordered pairs was higher.

The fact that ordering did not affect the perception of English vowels while having a significant effect on Swedish is of great interest, as it indicates that speakers' familiarity with a language may affect how ordering is used as a cue. As speakers are unfamiliar with Swedish vowels, they likely use additional cues in their environment to make judgments, such as vowel ordering. In this case, low vowels were chosen more often in Swedish high-low pairs than in the reverse order.

### 3.8.2 Vowel Pairs

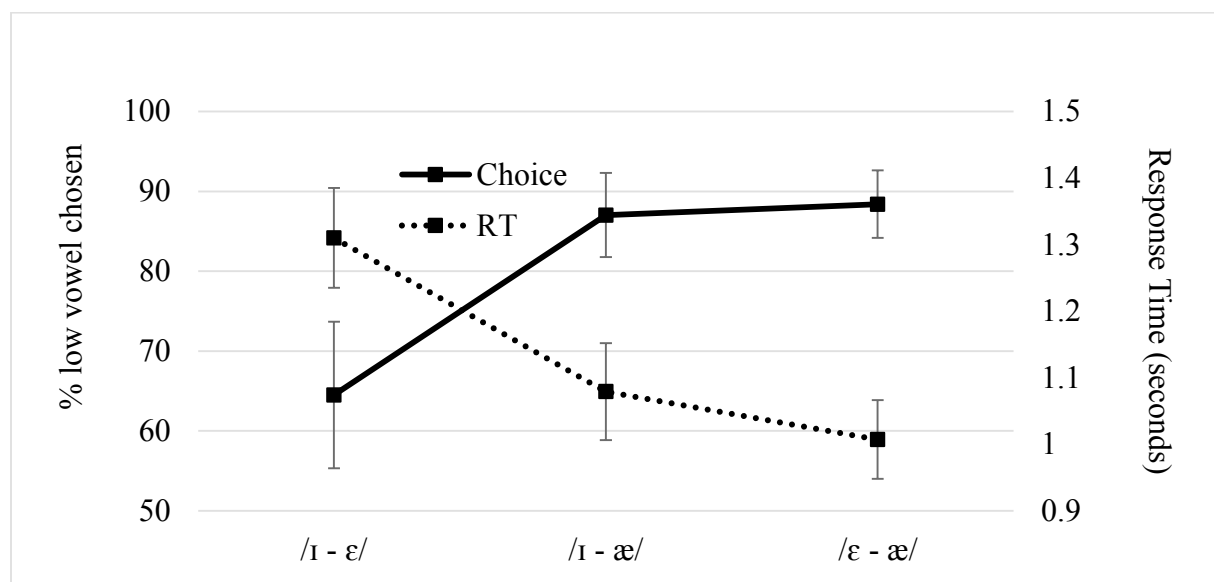
Across all conditions in English, low vowels in /ɪ - ε/ pairs were chosen as longer significantly less than in other pairs, and it took participants more time to make a choice. As shown in the preliminary measurements of vowels used as stimuli in *Section 3.4*, /ɪ/ and /ε/ are the least distinctive in terms of f1 height, and the closest in terms of average vowel length. Similar results are reported elsewhere in the literature, such as in Kessler, Treiman, and Mullennix (2002), Solé and Ohala, (2010), and Toivonen et al. (2014). Though /ɪ - æ/ pairs have the greatest height differences between vowels, low vowels were still chosen slightly less than in /ε - æ/ pairs, meaning that something other than height and duration differences affected duration perception.

Of the English vowels recorded for this study, combinations of /ε/ and /æ/ have the most similar f2 (backness). Speculatively, perhaps vowels are easiest to compare to one another when they have a similar backness. This is supported by Pfitzinger (2005), who found that an increase in backness (lowering of f2) is correlated slightly with vowels sounding higher, when contrasted with other vowels. Because /æ/ is further back than /ɪ/, perceptible differences between the two may be less than measurable differences. This would also explain the smallest perceived



differences between /ɪ - ε/, which also have a large f2 difference. Figure 9 illustrates similarities between how often and how quickly low vowels were chosen.

Figure 9: Response Time and Length Choice for English Vowels



As the pairs containing low vowels become lower, they get chosen as long more often. At the same time, they are chosen more quickly than high vowels.

In addition, Ryalls and Lieberman (1982) showed that American English speakers categorized the durations of lower vowels more accurately than with vowels in other conditions, as lower vowels offer a denser spectral sampling (assortment of vowel quality information) than higher vowels (p.1633). This can be elaborated on by Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, and Lacerda (1997), who explain that component frequencies in low vowels spoken to both adults and infants are more concentrated and compact (p.685). Lindblom (1986) also found low vowels to have a greater loudness density (p.26). However, because the above explanations are so varied in their analyses, further investigation is required in this area.

Little difference was found between /y - ø/ and /ø - ʊ/ in either order, apart from low vowels in the high-low vowel order being chosen significantly more often than when reversed.

Furthermore, response times were not significantly shorter when low vowels were chosen than when high vowels were chosen. If we were to consider the results from Pfitzinger (2005), this would come as a surprise, as the backness of each vowel in each pair is considerably different from one another. Thus speakers' unfamiliarity with the vowels must too be important. Flege, Bohn, and Jang (1997) indicated that inexperienced second-language (L2) learners of English are less sensitive to spectral cues (such as  $f_2$ ) than experienced L2 learners or native speakers. This potentially explains why native English speakers unfamiliar with Swedish are influenced by backness differences in English, but not in Swedish. This also provides further evidence for why ordering is so important in judging Swedish vowel length, and why participants choose significantly fewer low vowels when they are unable to use ordering preferences as a crutch.

### 3.8.3 Natural vs. Modified Duration

In both languages, participants' judgments of which vowel was longer did not change when duration was modified so that vowels were of an equal duration. As participants were unaware of a condition where vowels were modified to be the same length, identical durations were something they neither expected, nor had the opportunity of choosing during the task itself. After participants finished the experiment, some reported many of the pair words sounding identical to one another, and that they had just guessed the answer. However, the results confirm that at some level they made informed judgments.

Because of this, participants may have in fact been convinced by their previous lexical knowledge and biases that low vowels sound longer, even when they are measurably the same length as comparatively higher vowels. As native English speakers use duration to distinguish vowels (Gottfried and Beddor, 1988), participants in this study may have relied on previous

lexical knowledge to complete the task. Participants performed better across both duration conditions in English than in Swedish, and performed significantly worse in the low-high vowel Swedish condition, where they could not use vowel ordering to assist them.

A question remains, however: when durations were controlled to be the same length, did participants perform similarly to the natural length condition because they were consciously using previous vowel length knowledge, or because their perceptual biases “tricked” them into thinking the vowels were of different lengths? RT measures in this study offer some interesting possibilities.

In English, little RT change occurs between pairs with natural and modified durations for both /ɪ - ε / and /ɪ - æ / pairs, meaning that at least for these pairs, participants’ cognitive demand didn’t increase and was the same when vowels were controlled to be the same length. However, a more pronounced effect is noted for /ε - æ / pairs, where RT significantly increases when the durations of word pairs become modified. This again may be due to low vowels’ greater density of spectral information available to speakers (Ryalls & Lieberman, 1982), which then is partially removed when the vowels are modified for duration.

In Swedish normally ordered pairs, participants chose more low vowels as long in the modified condition than in the natural duration condition. However, in reverse-ordered pairs, the proportion of low vowels chosen is lower, and more low vowels are chosen in the natural duration condition. No difference in RT is shown across both pair orientations and modified/natural duration conditions. Speculatively, this difference in choice across reverse and normally ordered pairs with no RT differences potentially indicates that participants were less sure of their length choices due to their unfamiliarity with Swedish vowels. It should be noted that all

conditions mentioned here for Swedish still had low vowels chosen as longer significantly more than high vowels.

#### 3.8.4 Real vs. Nonsense Words

Participants made significantly more choices with low vowels in the nonsense word condition than in the real word condition, and had a significantly shorter RT with nonce words than with real words. This suggests either that semantic meaning in fact has an interfering effect on vowel judgment, or that removing this semanticity allows for increased accuracy and a lesser cognitive load. These findings are somewhat controversial, as tasks involving meaningful words have previously been shown in the literature to be more readily and easily performed than tasks with nonsense words. It should be noted that though this difference is *significant*, it is still only a 3.56% change, so whether or not these results are *meaningful* is up for debate.

Dollaghan, Biber, and Campbell (1992) found that children's memory for nonsense words was influenced by their memory of known words, and in this way the real and nonce words were embedded in one another. Interference occurred from accessing a meaningful prototype of the nonce word as a comparison in order to make a judgment. However, the results found here have the opposite results, so a further understanding of how semanticity interacts with other conditions in the experiment is necessary.

Though no ISI difference was observed in participants' choices, RT differences between real and nonce words that occurred in the Short ISI condition disappeared in the Long ISI condition. Furthermore, RT scores for both real and nonce words were shorter in the Long ISI condition than in the Short. This may support the findings in this thesis that semanticity interferes with vowel identification. As in pairs with a Short ISI more readily elicit surface-level

comparisons between words (Werker & Logan, 1985), the presence of semantic meaning could cause participants to attempt access their mental lexicons in addition to this surface-level processing, requiring added cognitive processing and time. As nonce words have no meaning attached to them, participants may not need to compare them to given entries in the lexicon, thus removing the additional load requirements and speeding up their choices. As a Long ISI is thought to correlate with participants comparing stimuli to similar forms in their mental lexicons, the interference of semanticity with real words found in the Short ISI condition remains puzzling.

In the long ISI, RT is similar for real and nonce pairs, indicating that a similar cognitive load is found for real and nonce words. Wang et al. (1976) found that participants perceived vowel stimuli in a consistent pattern, regardless of whether or not the forms they were presented with contained meaning, but no information on the time step between pair stimuli was discussed. It is thus possible that an ISI similar to a Long ISI in this study was used, yielding similar results. The literature does not seem to explain this interaction between semanticity and ISI, so future research in this area is crucial.

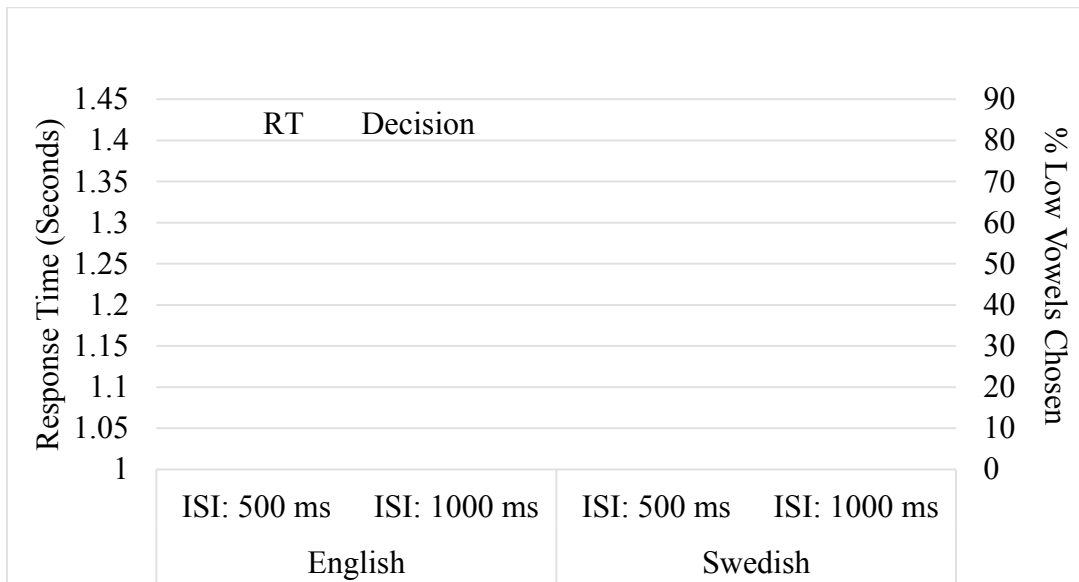
These findings go against the predictions outlined in this study's methods sections, as (comparing nonce words to infrequent words) the "less frequent"-like words have low vowels chosen more often and more quickly. However, the case that nonce words are similar to less frequent meaningful words is more than mere speculation, and the literature suggests that the way humans encode and remember (Majerus, Van der Linden, Mulder, Meulemans, & Peters, 2004) spoken information is affected by phonological frequencies (Vaden, Halpen, & Hickock, 2009). Because of this, the proximity of the nonce word's form to a real word may have an effect on how much interference occurs.

To test the proximity of nonce words to real words, nonce words were inputted into the Irvine Phonotactic Online Dictionary (IPhOD), Version 2.0 (Vaden et al., 2009). The IPhOD tests how many real-word neighbours a given nonce word has, which acts as a viable estimation for how closely particular nonce words resemble real words, and thus how separate from real words a given nonce word in actuality is. Real-word (phonotactic) neighbours are defined as words with semantic meaning which differ from the nonce word by one phoneme (speech sound) (Vaden et al., 2009). To support this notion, Vaden, Kuchinsky, Keren, Harris, Ahlstrom, Dubno, and Eckert (2011) suggest that a larger number of phonotactic neighbours can indirectly make nonce words more recognizable, by passively activating sub-lexical representations in the Left Inferior Frontal Gyrus (p.3570). As few other studies have looked at how the number of phonotactic neighbours affect intelligibility or similarity to real words, it is difficult to determine how the nonce words in this experiment compare to nonce words used in other experiments without performing a meta-analysis outside the scope of this thesis. Results for how many phonotactic neighbours are found for each nonce word are in *Appendix G*.

### 3.8.5 Short vs. Long ISI

Overall in the Choice data, there were no significant differences between short and long Inter-Stimulus Intervals for either English or Swedish. However, some interesting patterns can be found when compared to RT data, as shown below in Figure 10.

Figure 10: English and Swedish Length Choices, RT and ISI



Significant differences occurred in RT for English, suggesting that though there was little difference between ISIs in the *amount* low vowels were chosen over high vowels, participants were able to make a choice significantly *faster* when they were both familiar with the language, and the ISI was longer. Interestingly, this contradicts previous findings in Werker and Logan (1985), that associate faster processing speeds with Short ISIs that elicit auditory surface level processing, indicating a lesser processing demand. However, it should be noted that semanticity was not a factor in their experiment, which used nonsense speech-syllables with no codas.

For English nonce stimuli, (non-significantly) shorter RTs were in the Short ISI, yielding similar results to the previous findings. RTs were significantly longer when semanticity was introduced in the real-word condition, presumably because participants attempted to match the stimulus to entries in their mental lexicon (Goldinger, 1998), even though only auditory-level processing was elicited. They thus had to perform two different types of processing simultaneously, increasing their response time.

A longer ISI caused participants to access phonetic-level processing (Werker & Logan, 1985) for both real and nonce conditions, whereby they accessed their grammars to make vowel length judgments. The presence of meaning also caused them to involuntarily access these mental representations in a potentially similar way, meaning that they may have had to perform the one task of lexical access, and not have had to perform acoustic comparisons as well.

No Choice and RT differences for Swedish indicate that a lack of familiarity may influence how participants judge vowel length. As speakers are presumably unfamiliar with Swedish, they have no previous Swedish entries in their mental lexicons, and are thus unable to directly link the sounds they hear with previous knowledge. Liberman, Harris, Hoffman, and Griffith (1957) stipulate that individuals encountering sounds in an unfamiliar language tend to categorize these sounds according to what they most closely resemble in the speaker's knowledge (p.358), as they are not as sensitive to vowel quality differences (Flege, Bohn, & Jang, 1997). As speakers do not understand the meaning of the words, they may revert to a form of processing similar to acoustic processing, which would explain the identical response time results across ISI.

### 3.8.6 Limitations

As this study is a first step in experimental procedure regarding forced-choice multi-factor vowel perception tasks, there is still considerable room for improvement. Volunteered subjective reports suggest that the task was somewhat repetitive and boring, but this is somewhat unavoidable to this type of task. However, some participants reported being able to tell that some words were indeed controlled to be exactly the same length.



Because 40 experimental stimuli were selected for each block from a larger database for the purpose of shortening the experiment, some stimuli were not included. Because of this and other potential errors in incorporating stimuli in the experiment, some sub-conditions (e.g. the /e/ - /o/ pair in the Swedish 500-millisecond ISI modified duration Choice condition) were not accounted for in the results. Any future studies looking at these methods should ensure that these errors are minimized, so that all conditions can be entirely present in the experiment. Ensuring the presence of all conditions would also further increase the results' reliability, as statistical tests would use more balanced data.

The English stimuli consisted of pair words that only contained voiced codas, to ensure that coda voicing was not an issue. However, it would be interesting to further investigate voicing on vowel length perception, as voiced consonants are shown to lengthen any preceding vowels (House & Fairbanks, 1953; Klatt, 1976; Kluender, Diehl, & Wright, 1988).

As vowels of interest were all contained in words, there are likely environmental factors (such as onset voicing) which affected vowel perception. Such effects should be minimal, as the stimuli were presented as minimal pairs, but this could be completely controlled for in a future experiment by presenting vowels in isolation. However, this would present an additional issue as the task itself would become more transparent to participants, meaning they may better understand the nature of the task, which could potentially alter the results. Furthermore, semanticity could not be tested as vowels in isolation have no inherent meaning in English.

Finally, given more time and financial resources, experimental validity could be increased by improving the quality of technology used to conduct the experiment. Though the task was successfully implemented using the Acer Aspire E1 laptop, the computer itself was not designed for highly time-sensitive perception experiments, which resulted in occasional minor

glitches in processing speed and stimulus presentation. Though this likely had a negligible effect on response time information, the full impact of potential glitches is unknown.

#### 4.0 General Discussion

Studies 1 and 2 in this thesis suggest that not only are low vowels *measurably* longer than high vowels, but that they also *sound* longer. Study 1 indicates that DHC is controlled, and likely not entirely mechanical. In the literature, using length to distinguish vowels from one another is observed to be an essential trait of both early language acquisition (Bohn & Polka, 2001) and vowel discrimination in adults (Gottfried & Beddor, 1988). Duration is also important in improving the intelligibility of vowels when the presence of noise causes spectral information to be degraded (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995, p.303). If this is the case, then some cognitive biases affecting perception seem to be a viable alternative to a purely physiological explanation to DHC.

Perceptual evidence from Study 2 shows that speakers judge pairs with only non-high vowels more quickly and reliably than high vowels, suggesting that vowels differ in terms of height in the way they are perceived. This may be because low vowels tend to have a greater amount of vowel quality information (Ryalls & Lieberman, 1982, p. 1633), which in turn affects speakers' preferences. If this is indeed the case, then low vowels may be both produced and perceived as longer because they carry more information. Low vowels are longer because of the comparatively high amount of information they carry. Speakers may begin experiencing this inherent "longness" in pre-linguistic infancy when they require vowel length to discern between vowels (Bohn & Polka, 2001, p.512; Lehiste 1970, p.131), leading them to develop unconscious associations between low vowels and increased duration that is observed in adulthood (Gottfried

& Beddor, 1988, p.58). In this case, speakers may have a particular disposition to accurately perceiving vowel length (and develop associations after language learning occurs), but do not attach longer durations to lower vowels because of some innate drive. Perception across languages was tested in Study 2 in an attempt to understand to what extent the perception of DHC is learned (or at least not innate).

Speakers are able to generalize their duration judgments to languages with sound systems they are unfamiliar with, with considerable accuracy. This being said, their judgments are still less reliable and take a longer time to be made in unfamiliar languages, suggesting that judgments are at least partially learned. As shown with ordering effects in Study 2, listeners of an unfamiliar language will attempt to use their intuitions and any additional cues available to them to make a choice. Intuitions such as preferring reduplicative pairs with high vowels followed by low vowels (e.g. “zig-zag”) are found across languages (Polka & Bohn, 2003), and were demonstrated to play an integral role in determining vowel length in this thesis.

As participants selected low vowels as longer significantly more often than high vowels, they were likely sensitive to the differences in vowel length between high and low vowels even though they had never heard the vowel sounds before. Because they selected Swedish low vowels to a lesser extent on average (when controlled for ordering effects), this suggests that at least some of the judgments listeners make is based on prior learning. This is supported by Flege, Bohn, and Jang (1997), whereby unexperienced speakers’ abilities to perceive differences in a specific second language’s vowel durations are smaller and more constrained to rules in their native language than familiar speakers (p.2548). Further evidence from Study 2 indicates that this learning may be in the form of associations speakers develop from hearing longer low vowels throughout their lifetime.

A high choice rate and low response time for word pairs with two non-high vowels in a familiar language indicates that length judgments can most easily be made for low vowels, meaning that low vowels (which carry the densest spectral information) are longest to produce, and take the least amount of time to perceive (though they are still also perceived as being the longest). Though low vowels contain a greater density of spectral information, they seem to be more easily perceived than high vowels. Speculatively, the added length could allow listeners more time to properly perceive the more informative vowels, while this greater density of vowel information available allows for greater accuracy. This sort of compensatory mechanism could have become phonologized across languages, but to date this theory has not been explored.

#### 4.1 Future Research

These findings provide an alternative explanation for why the DHC is seemingly specified in a language's grammar, yet so prevalent cross-linguistically. As opposed to being an entirely physiological behaviour specified by factors such as jaw length (Lehiste, 1970, p.19), a universal preference for long low vowels may be because they are perceived as longer, perhaps due to low vowels inherently containing more spectral information (Ryalls & Lieberman, 1982; Kuhl et al., 1997), which results in vowels being perceived as longer. However, little research has investigated why low vowels carry more information audible to humans than high vowels, and most studies reporting this phenomenon only report its existence.

This increase in spectral information and compactness may be determined by how the human auditory pitch range interacts with  $f_1$  frequencies. In this scenario, low vowels (which have a high  $f_1$ ) would be more audible and carry more information accessible to humans as they exist in the hearing-range that is most comfortably and reliably perceived. No studies have been

found by this author in the literature that report relationships between human hearing range and the f1 frequency distributions of vowels.

Meaning behind words seem to interfere with length perception when acoustic processing is elicited as it causes listeners to access their mental lexicons while only enough time is provided for a surface-level comparison to be made. Though this does not appear to affect speakers' abilities to judge length, it does cause the task itself to become more difficult, causing speakers to spend more time making the choice. As this is mere speculation in the face of a literature that suggests otherwise, additional research should focus on better understanding what happens when listeners are required to process semantic meaning at a surface, acoustic level.

Furthermore, a better understanding of how each of the perceptual variables mentioned above affect how speech sounds are produced is necessary. For instance, though semanticity has been shown to affect certain aspects of vowel length perception, how does it affect *production*? Answering questions such as these would be crucial in obtaining a better understanding of how exactly the perception and production of DHC are linked together.

Overall, the production and perception of DHC appear to relate to each other through both learned associations and as duration as an essential property of vowels. Though this is a necessary step in better understanding how duration is employed in speech, extensive research still needs to investigate further intricacies in the relationship between duration perception and production, specifically with regards to how much of the DHC is learned, and how much of it is specified by physiology, perception, and how both of these interact with the physics of what vowel information is available to listeners.

## 5.0 Conclusion

From two experiments investigating the production and perception of vowels, a conclusion can be drawn that high vowels are both measurably and perceivably shorter than low vowels. This duration-height correlation is shown to typically be specified in a given language's grammar, but occurs in similar ways cross-linguistically due to how listeners perceive vowels of different heights. Vowel length judgments are learned to the extent that they can be extended to word pairs with vowels that have been controlled to be the same length, at least when in a familiar language. Though listeners perceive low vowels in an unfamiliar language as significantly longer than high vowels, they still do so less of the time than in a language they have native proficiency in, and rely on external cues such as ordering to make a choice. This suggests that some – but not all – of their perceptual abilities come from learned associations in their native language, which they then attempt to extend to other languages. This behaviour has been observed elsewhere in the literature (Flege, Bohn, & Jang, 1997, p.2548).

In addition, listeners make vowel judgments more quickly when the processing demands of the task match the type of processing that has been elicited. In this way, listeners attempting to access their mental lexicons when they only have time to make surface-level acoustic comparisons take longer to make a length choice, than when the condition matches their processing level or they encounter nonsense words with no meaning attached. However, little is still known about how this would manifest itself in speech production with relation to vowel duration. In this thesis, the production study indicated that DHC is controlled, and not mechanical in nature. The second study then investigated the extent to which these seemingly controlled origins are perceptual in nature. However, further research needs to analyze how these

perceptual findings relate back to production, to better understand the relationship between the production and perception of vowel height and duration.

## 6.0 References

- Abramson, A. S. (1962). *The vowels and tones of standard Thai: Acoustical measurements and experiments* (Vol. 20). Indiana University.
- Äimä, F. (1918). Phonetik und Lautlehre des Inarilappischen I-II. *Memoires de la Societé Finno-Ougrienne XIII-XLIII. Helsinki.*
- Bauer, L., & Warren, P. (2008). New Zealand English: phonology. *Varieties of English*, 3, 39-63.
- Bergmann, A., Hall, C.K., Ross, S.M. (2007) *Language Files*, The Ohio State University Press, Columbus, 10.
- Bernstein, B. (2013). Clases sociales, lenguaje y socializacion. *Infoamerica - Iberoamerican Communication Review*, 9. Retrieved November 23, 2014, from [http://www.infoamerica.org/documentos\\_pdf/bernstein05.pdf](http://www.infoamerica.org/documentos_pdf/bernstein05.pdf)
- Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program]. Version 5.4.04, retrieved 21 January 2015 from <http://www.praat.org/>
- Bohn, O. S., & Polka, L. (2001). Target spectral, dynamic spectral, and duration cues in infant perception of German vowels. *The Journal of the Acoustical Society of America*, 110(1), 504-515.
- Bradlow, A. R., Torretta, G. M., and Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics, *Speech Commun.* 20, 255–272.
- Brigner, W. L. (1988). Perceived duration as a function of pitch. *Perceptual and motor skills*, 67(1), 301-302.
- Brunken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 53-61.



- Bye, P., Sagulin, E., & Toivonen, I. (2010). Phonetic duration, phonological quantity and prosodic structure in Inari Saami. *Phonetica*, 66(4), 199-221.
- Byrd, D. (1992) Sex, dialects, and reduction. *Proceedings ICSLP 92: 1992 International Conference on Spoken Language Processing*, 1, 827-830.
- Calvin, W. H. (1983). A stone's throw and its launch window: timing precision and its implications for language and hominid brains. *Journal of Theoretical Biology*, 104(1), 121-135.
- Catford, J. (1977). Co-articulation and modified articulation. In *Fundamental Problems in Phonetics* (Vol. 1, pp. 188-209). The University of California: Indiana University Press.
- Chomsky, N. (1972). *Language and mind* (p. 100). New York: Harcourt Brace Jovanovich.
- Cooper, W. E., & Klouda, G. V. (1995). The psychological basis of syntactic iconicity. *Syntactic Iconicity and Linguistic Freezes: The Human Dimension*, 9, 335.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, 11(6), 671-684.
- Dollaghan, C., Biber, M., & Campbell, T. (1992, June). Words within nonwords: Constituent syllable effects in a measure of phonological working memory. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.
- Eilers, R. E., Bull, D. H., Oller, D. K., & Lewis, D. C. (1984). The discrimination of vowel duration by infants. *The Journal of the Acoustical Society of America*, 75(4), 1213-1218.
- Elert, C.C. (1964). Phonologic studies of Quantity in Swedish. *Uppsala*.
- Engstrand, O., & Krull, D. (1994). Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica*, 51(1-3), 80-91.

- Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners. *The Journal of the Acoustical Society of America*, 116(4), 2365-2373.
- Fischer-Jørgensen, E. (1964). Paul Diderichsen: Helhed og struktur *GEC Gads Forlag*, 11-19.
- Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of phonetics*, 25(4), 437-470.
- Fujisaki, H., & Kawashima, T. (1969). On the modes and mechanisms of speech perception. *Annual Report of the Engineering Research Institute*, 28, 67-73.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, 105(2), 251.
- Gottfried, T. L., & Beddor, P. S. (1988). Perception of temporal and spectral information in French vowels. *Language and Speech*, 31(1), 57-75.
- Gussenhoven, C. (2007). A vowel height split explained: Compensatory listening and speaker control. *Laboratory Phonology*, 9, 145-172.
- Gussenhoven, C., & Zhou, W. (2013). Revisiting pitch slope and height effects on perceived duration. In *INTERSPEECH* (pp. 1365-1369).
- Hardison, D. M., & Saigo, M. M. (2010). Development of perception of second language Japanese geminates: Role of duration, sonority, and segmentation strategy. *Applied Psycholinguistics*, 31(01), 81-99.
- Heffner, R. S. (1937). Notes on the length of vowels. *American Speech*, 128-134.
- Hoemeke, K. A., & Diehl, R. L. (1994). Perception of vowel height: The role of F1-F0 distance. *The Journal of the Acoustical Society of America*, 96(2), 661-674.

- House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America*, 25, 105.
- House, A. S. (1961). On vowel duration in English. *The Journal of the Acoustical Society of America*, 33(9), 1174-1178.
- Hualde, J. I. (2005). *The sounds of Spanish*. New York: Cambridge University Press.
- Jacewicz, E., Fox, R. A., & Salmons, J. (2006). Prosodic prominence effects on vowels in chain shifts. *Language Variation and Change*, 18(03), 285-316.
- Jakobson, R. (1968). Foundation of the Structural Laws. *Child language, aphasia, and phonological universals* (pp. 67-94). The Hague: Mouton.
- Järvikivi, J., Vainio, M., & Aalto, D. (2010). Real-time correlates of phonological quantity reveal unity of tonal and non-tonal languages. *PloS one*, 5(9), e12603.
- Jensen, A. R. (1993). Why is reaction time correlated with psychometric g? *Current Directions in Psychological Science*, 53-56.
- Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam: Elsevier.
- Johnson, K., & Martin, J. (2000). Acoustic vowel reduction in Creek: Effects of distinctive length and position in the word. *Phonetica*, 58(1-2), 81-102.
- Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language*, 47(1), 145-171.
- Kingston, J., Kawahara, S., Chambless, D., Mash, D., & Brenner-Alsop, E. (2009). Contextual effects on the perception of duration. *Journal of Phonetics*, 37(3), 297-320.

- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America*, 59(5), 1208-1221.
- Kluender, K. R., Diehl, R. L., & Wright, B. A. (1988). Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics*.
- Kondaurova, M. V., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *The Journal of the Acoustical Society of America*, 124(6), 3959-3971.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., ... & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277(5326), 684-686.
- Labov, W. (1994). *Principles of linguistic change. 1: Internal factors*. Oxford: Blackwell.
- Lee, J. Y. (2008). Perception of English High Vowels: Duration as a Cue by Korean Speakers of English.
- Leinonen, T. N. (2010). *An acoustic analysis of vowel pronunciation in Swedish dialects*. Rijksuniv.
- Lehiste, I. (1971). Temporal organization of spoken language. *Form and Substance*, 159-169.
- Lehiste, I. (1982). Some phonetic characteristics of discourse\*. *Studia Linguistica*, 36(2), 117-130.
- Lehnert-LeHouillier, H. (2010). A cross-linguistic investigation of cues to vowel length perception. *Journal of Phonetics*, 38(3), 472-482.

- Lennes, Mietta . "SpeCT - The Speech Corpus Toolkit for Praat." (*former title: Mietta's scripts for the Praat program*). helsinki.fi, 20 Jan. 2011. Web. 23 Sept. 2014.  
<<http://www.helsinki.fi/~lennes/praat-scripts/#calculate>>.
- Liberman AM, Harris KS, Hoffman HS, Griffith BC. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*. 1957; 54:358–368. [PubMed: 13481283]
- Lindblom, B. (1967). Vowel duration and a model of lip mandible coordination. *Speech Transmission Laboratory Quarterly Progress and Status Report*, 4(1967), 1-29.
- Lindblom, B. (1986). Phonetic universals in vowel systems. *Experimental phonology*, 13-44.
- Lipski, J. M. (2008). *Varieties of Spanish in the United States*. Georgetown University Press. (pp. 199- 201).
- Lisker, L. (1974). Is it "VOT" or a first-formant transition feature detector? *Paper presented at the annual meeting of the American Association of Phonetic Sciences*, St. Louis, Missouri.
- Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network. *Journal of Memory and Language*, 51(2), 297-306.
- Mateer, C., & Kimura, D. (1977). Impairment of nonverbal oral movements in aphasia. *Brain and Language*, 4(2), 262-276.
- McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of phonetics*, 30(2), 229-258.

- McLeod, S. (2007). Levels of Processing. *Simply Psychology*. Retrieved from:  
<http://www.simplypsychology.org/levelsofprocessing.html>
- Meador, K. J., Ray, P. G., Echauz, J. R., Loring, D. W., & Vachtsevanos, G. J. (2002). Gamma coherence and conscious perception. *Neurology*, *59*(6), 847-854.
- Meister, E., & Werner, S. (2009). Duration affects vowel perception in Estonian and Finnish. *Linguistica Uralica*, *(3)*, 161-177.
- Minagawa-Kawai, Y., Mori, K., & Sato, Y. (2005). Different brain strategies underlie the categorical perception of foreign and native phonemes. *Journal of cognitive neuroscience*, *17*(9), 1376-1385.
- Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and speech*, *38*(3), 289-306.
- Ojemann, G., & Mateer, C. (1979). Human language cortex: localization of memory, syntax, and sequential motor-phoneme identification systems. *Science*, *205*(4413), 1401-1403.
- Payne, D. (2008, August 15). Contrastive Vowels. *Contrastive Vowels*. Retrieved September 17, 2014, from <http://darkwing.uoregon.edu/~dlpayne/maasai/ATR.htm>
- Pater, J. (2000). Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints. *Phonology*, *17*(02), 237-274.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175-184.
- Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America*, *32*, 693.
- Peirce, J.W. (2007). PsychoPy - Psychophysics software in Python. *J Neurosci Methods*, *162* (1-2), 8-13.

- Pfitzinger, H. R. (2005). Towards functional modelling of relationships between the acoustics and perception of vowels. *ZAS papers in Linguistics*, 40, 133-144.
- Pinker, Steven. 1994. *The Language Instinct*. New York: HarperPerennial.
- Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, 13(2), 253-260.
- Pisoni, D. B. (1975). Auditory short-term memory and vowel perception. *Memory & Cognition*, 3(1), 7-18.
- Polka, L., & Bohn, O. S. (2003). Asymmetries in vowel perception. *Speech Communication*, 41(1), 221-231.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 421.
- Price, C. J., Veltman, D. J., Ashburner, J., Josephs, O., & Friston, K. J. (1999). The critical relationship between the timing of stimulus presentation and data acquisition in blocked designs with fMRI. *NeuroImage*, 10(1), 36-44.
- Raphael, L. J. (1972). Preceding Vowel Duration as a Cue to the Perception of the Voicing Characteristic of Word-Final Consonants in American English. *The Journal of the Acoustical Society of America*, 51(4B), 1296-1303.
- Ryalls, J. H., & Lieberman, P. (1982). Fundamental frequency and vowel perception. *The Journal of the Acoustical Society of America*, 72(5), 1631-1634.
- Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D., & Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science*, 326(5951), 445-449.

- Savin, H. B. (1963). Word-Frequency Effect and Errors in the Perception of Speech. *The Journal of the Acoustical Society of America*, 35(2), 200-206.
- Scharf, B. (1962). 'Loudness summation and spectrum shape. *The Journal of the Acoustical Society of America*, 34, 228–233.
- Schubotz, R. I., Friederici, A. D., & Yves von Cramon, D. (2000). Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *Neuroimage*, 11(1), 1-12.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303-304
- Solé, M. J. (2007). Controlled and mechanical properties in speech. *Experimental approaches to phonology*, 302-321.
- Solé, M. J., & Ohala, J. J. (2010). What is and what is not under the control of the speaker: Intrinsic vowel duration. *Laboratory phonology*, 10.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta psychologica*, 30, 276-315.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6), 643.
- Styler, W. (2013). Creating and manipulating sound files in PRAAT. In Using PRAAT for linguistic research (Vol. 1.3.5, pp. 32-35). Boulder, Colorado: University Colorado at Boulder Phonetics Lab.
- Swoboda, P. J., Kass, J., Morse, P. A., & Leavitt, L. A. (1978). Memory factors in vowel discrimination of normal and at-risk infants. *Child Development*, 332-339.



- Tauberer, J., & Evanini, K. (2009). Intrinsic vowel duration and the post-vocalic voicing effect: some evidence from dialects of North American English. In *INTERSPEECH* (pp. 2211-2214).
- Toivonen, I., Blumenfeld, L., Gormley, A., Hoiting, L., Logan, J., Ramlakhan, N., & Stone, A. (2014). Vowel Height and Duration. *West Coast Conference on Formal Linguistics (WCCFL)*, 32.
- Tomás, T. N. (1916). *Cantidad de las vocales acentuadas*. Imprenta de los Sucesores de Hernando.
- Umeda, N. (1975). Vowel duration in American English. *The Journal of the Acoustical Society of America*, 58(2), 434-445.
- Vaden, K.I., Halpin, H.R., Hickok, G.S. (2009). Irvine Phonotactic Online Dictionary, Version 2.0. [Data file]. Available from <http://www.iphod.com>.
- Vaden, K. I., Kuchinsky, S. E., Keren, N. I., Harris, K. C., Ahlstrom, J. B., Dubno, J. R., & Eckert, M. A. (2011). Inferior frontal sensitivity to common speech sounds is amplified by increasing word intelligibility. *Neuropsychologia*, 49(13), 3563-3572.
- Van Dommelen, W. (1982). A contrastive investigation of vowel duration in German and Dutch. *Phonetica*, 39(1), 23-35.
- Viceni, C. (2009a). Flat-intonation Scaler. Retrieved January 21, 2015, from <http://www.linguistics.ucla.edu/faciliti/facilities/acoustic/praat.html#filemanagement>
- Viceni, C. (2009b). Intensity Scaler. Retrieved January 21, 2015, from <http://www.linguistics.ucla.edu/faciliti/facilities/acoustic/praat.html#filemanagement>
- Wang, W. S. Y., Lehiste, I., Chuang, C. K., & Darnovsky, N. (1976). Perception of vowel duration. *The Journal of the Acoustical Society of America*, 60(S1), S92-S92.

- Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, 37(1), 35-44.
- Ylinen, S., Shestakova, A., Alku, P., & Huotilainen, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech*, 48(3), 313-338.
- Ylinen, S., Shestakova, A., Huotilainen, M., Alku, P., & Näätänen, R. (2006). Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study. *Brain Research*, 1072(1), 175-185.
- Zimmerman, S. A., & Sapon, S. M. (1958). Note on Vowel Duration Seen Cross-Linguistically. *The Journal of the Acoustical Society of America*, 30(2), 152-153.

*Appendix A*

List used for Study 1 on Chilean Spanish. Each list consists of the same 21 words, but in a different randomized order.

<b>List 1</b>	<b>List 2</b>	<b>List 3</b>	<b>List 4</b>	<b>List 5</b>
misa	lesa	biso	baso	lesa
baso	biso	baso	saco	lizo
saco	Lisa	paso	lesa	paso
mesa	misa	saco	masa	mesa
seco	paso	lasa	paso	masa
paso	saco	lesa	Lisa	menos
Lisa	baso	mino	beso	mino
mino	piso	Lisa	biso	cida
lazo	mesa	misa	menos	piso
mano	beso	ceda	lazo	saco
lesa	cida	mesa	mano	peso
beso	lizo	masa	mino	biso
cida	masa	lazo	seco	misa
masa	mino	lizo	misa	lazo
piso	ceda	beso	piso	beso
menos	peso	menos	lizo	lasa
lizo	lazo	mano	peso	baso
peso	seco	cida	cida	ceda
ceda	mano	piso	ceda	seco
biso	lasa	seco	lasa	Lisa
lasa	menos	peso	mesa	mano

*Appendix A*

<b>List 6</b>	<b>List 7</b>	<b>List 8</b>	<b>List 9</b>	<b>List 10</b>
ceda	mano	beso	saco	masa
Lisa	beso	cida	mano	lasa
beso	saco	masa	piso	peso
lesa	lazo	mano	peso	ceda
peso	menos	mesa	biso	paso
mano	lasa	lizo	lizo	lizo
biso	Lisa	biso	paso	menos
lasa	biso	seco	menos	seco
lizo	lesa	ceda	baso	misa
piso	piso	menos	Lisa	mesa
seco	ceda	baso	masa	beso
masa	seco	piso	mesa	lazo
paso	cida	lesa	seco	piso
cida	paso	saco	misa	saco
saco	baso	mino	beso	mino
menos	lizo	lazo	lazo	baso
mesa	peso	Lisa	ceda	lesa
mino	mino	misa	lesa	cida
misa	masa	lasa	lasa	Lisa
baso	mesa	peso	mino	biso
lazo	misa	paso	cida	mano

*Appendix B*

The following Appendix shows T-Tests for *between-category* distinctions in Study 1.

T-Test for the **duration** [i] - [e] ; [e] - [a] (Respectively)

Speaker	t	df	p-value	95% confidence		Sig. <0.05
PPT1	-3.794	133	<.000	-33.3221	-10.4852	**
PPT2	-3.799	138	<.000	-25.6227	-8.0127	**
PPT3	-5.354	137	<.000	-21.3364	-8.8264	**
PPT4	-3.253	132	<.000	-12.8216	-3.1247	**
PPT5	-6.195	124	<.000	-28.8216	-14.8633	**
PPT6	-3.711	138	<.000	-13.2612	-4.0415	**
PPT1	-1.779	135	0.078	-21.0019	1.1134	
PPT2	-5.605	137	<.000	-39.3467	-18.8233	**
PPT3	-2.245	136	0.026	-13.7696	-0.8734	*
PPT4	-3.985	136	<.000	-18.9179	-6.369	**
PPT5	-5.889	124	<.000	-30.6639	-15.2373	**
PPT6	-7.635	138	<.000	-24.7612	-14.5746	**

*Appendix B*

T-Test for the **f1** [i]- [e] ; [e] - [a] (Respectively)

Speaker	t	df	p-value	95% confidence		Sig. <0.05
PPT1	-11.184	133	<.000	-138.0680	-96.5715	**
PPT2	-16.02	138	<.000	-172.4894	-134.5879	**
PPT3	-15.017	137	<.000	-157.0815	-120.5267	**
PPT4	-15.184	132	<.000	-115.0575	-88.5349	**
PPT5	-9.306	124	<.000	-125.7656	-81.6511	**
PPT6	-12.261	138	<.000	-98.2128	-70.9347	**
PPT1	-26.327	135	<.000	-371.8434	-319.8810	**
PPT2	-22.573	137	<.000	-365.7513	-306.8324	**
PPT3	-32.122	136	<.000	-287.9184	-254.5235	**
PPT4	-16.187	136	<.000	-176.7406	-138.2568	**
PPT5	-21.447	124	<.000	-247.0358	-205.2903	**
PPT6	-13.326	138	<.000	-168.5846	-137.2565	**

*Appendix C*

Pearson's R for the Interaction between Duration and F1

***Correlations for [a]***

	<b>R Value</b>	<b>Alpha Score</b>	<b>Significance</b>	<b>Type</b>
PPT1	-0.322	0.007	**	Negative
PPT2	-0.076	0.531		Negative
PPT3	-0.525	<0.000	**	Negative
PPT4	0.285	0.016	*	Positive
PPT5	0.099	0.441		Positive
PPT6	0.058	0.636		Positive

***Correlations for [e]***

<b>Speaker</b>	<b>R Value</b>	<b>Alpha Score</b>	<b>Significance</b>	<b>Type</b>
PPT1	-0.171	0.164		Negative
PPT2	-0.083	0.498		Negative
PPT3	-0.498	<0.000	**	Negative
PPT4	-0.23	0.061		Negative
PPT5	-0.104	0.419		Negative
PPT6	0.079	0.517		Positive

***Correlations for [i]***

<b>Speaker</b>	<b>R Value</b>	<b>Alpha Score</b>	<b>Significance</b>	<b>Type</b>
PPT1	0.028	0.825		Positive
PPT2	-0.154	0.199		Negative
PPT3	0.091	0.451		Positive
PPT4	-0.174	0.159		Negative
PPT5	0.016	0.902		Positive
PPT6	-0.031	0.796		Negative

*Appendix D*

Stimulus list for Study 2 involving speaker perceptions in English.

Real Word Pairs		Nonce Word Pairs	
lid	led	dib	deb
lid	lad	dib	dab
led	lad	deb	dab
bid	bed	gid	ged
bid	bad	gid	gad
bed	bad	ged	gad
big	beg	kib	keb
big	bag	kib	kab
beg	bag	keb	kab
		mib	meb
		mib	mab
		meb	mab



*Appendix D*

Tentative stimulus list for Study 2 involving speaker perceptions in Swedish. NOTE: To match the length of the list for English stimuli, these minimal pairs will be repeated.

Real Word Pairs		Nonce Word Pairs	
gnytt	gnott	drytt	drott
grytt	grott	mytt	mott
rytt	rott	fytt	fott
bytt	bott	jytt	jott
knytt	gnott	klytt	klott
brytt	trott	gytt	gott

## *Appendix E*

This appendix contains the stimuli used in the perception experiment in Study 2. The practice block has 10 items, all 4 other blocks have 40. The items in the lists below occur as follows:

- Stimulus 1: The first word in the minimal pair.
- Stimulus 2: The second word in the minimal pair.
  - /m/ is placed at the beginning of either stimulus name if the duration has been manipulated
- ISI: The time between both pair words (Inter-Stimulus Interval) in milliseconds. An ISI of “1” refers to 500 milliseconds, and an ISI of “2” refers to 1000 milliseconds, or 1 second.
- Real/ Nonce: Whether the word has semantic meaning or not
- Manipulated?: Whether or not the word retains its natural duration, or has been manipulated.
  - naturaldur: natural duration intact
  - manipdur: duration of first word controlled to be identical to the second
- Orientation: Orientation is normal if Stimulus 1 is followed by Stimulus 2  
Orientation is reverse if Stimulus 2 is followed by Stimulus 1
- Language: The word pair’s language, either English or Swedish

The stimulus list for the Practice Trial Block is as follows:

Stimulus 1	Stimulus 2	ISI	Real/ Nonce	Manipulated?	Orientation	Language
mib.wav	mab.wav	1	nonce	naturaldur	normal	English
bid.wav	bed.wav	1	real	naturaldur	normal	English
mbeg.wav	mbig.wav	1	real	manipdur	reverse	English
mmott.wav	mmutt.wav	1	N/A	manipdur	reverse	Swedish
mdr6tt.wav	mdrytt.wav	1	N/A	manipdur	reverse	Swedish
led.wav	lid.wav	1	nonce	naturaldur	reverse	English
mkl6tt.wav	mklytt.wav	1	N/A	manipdur	reverse	Swedish
hytt.wav	f6tt.wav	1	N/A	naturaldur	normal	Swedish
n6tt.wav	nytt.wav	1	N/A	naturaldur	reverse	Swedish

mott.wav	mutt.wav	1	N/A	naturaldur	reverse	Swedish
----------	----------	---	-----	------------	---------	---------

*Appendix E*

The stimulus list for the Trial Block 1 is as follows, with an ISI of 500 ms.:

Stimulus 1	Stimulus 2	ISI	Real/Nonce	Manipulated?	Orientation	Language
sytt.wav	s6tt.wav	1	N/A	naturaldur	normal	Swedish
mt6tt.wav	mdytt.wav	1	N/A	manipdur	reverse	Swedish
t6tt.wav	dytt.wav	1	N/A	naturaldur	reverse	Swedish
mled.wav	mlid.wav	1	nonce	manipdur	reverse	English
meb.wav	mab.wav	1	nonce	naturaldur	normal	English
ged.wav	gad.wav	1	nonce	naturaldur	normal	English
mglott.wav	mplutt.wav	1	N/A	manipdur	reverse	Swedish
mmab.wav	mmeb.wav	1	nonce	manipdur	reverse	English
mf6tt.wav	mhytt.wav	1	N/A	manipdur	reverse	Swedish
msh6tt.wav	mshytt.wav	1	N/A	manipdur	reverse	Swedish
mgid.wav	mgad.wav	1	nonce	manipdur	normal	English
mmeb.wav	mmab.wav	1	nonce	manipdur	normal	English
mbid.wav	mbed.wav	1	real	manipdur	normal	English
mgad.wav	mgid.wav	1	nonce	manipdur	reverse	English
mmeb.wav	mmib.wav	1	nonce	manipdur	reverse	English
mdib.wav	mdab.wav	1	nonce	manipdur	normal	English
bad.wav	bed.wav	1	real	naturaldur	reverse	English
mgid.wav	mged.wav	1	nonce	manipdur	normal	English
mglytt.wav	mkludd.wav	1	N/A	manipdur	normal	Swedish
mbag.wav	mbig.wav	1	real	manipdur	reverse	English
keb.wav	kib.wav	1	nonce	naturaldur	reverse	English
ludd.wav	lodd.wav	1	N/A	naturaldur	normal	Swedish
mdrott.wav	mrutt.wav	1	N/A	manipdur	reverse	Swedish

mkeb.wav	mkib.wav	1	nonce	manipdur	reverse	English
mmott.wav	mmutt.wav	1	N/A	manipdur	reverse	Swedish
beg.wav	big.wav	1	real	naturaldur	reverse	English
trytt.wav	tr6tt.wav	1	N/A	naturaldur	normal	Swedish
kib.wav	kab.wav	1	nonce	naturaldur	normal	English
mlott.wav	mlutt.wav	1	N/A	manipdur	reverse	Swedish
big.wav	bag.wav	1	real	naturaldur	normal	English
mib.wav	meb.wav	1	nonce	naturaldur	normal	English
n6tt.wav	nytt.wav	1	N/A	naturaldur	reverse	Swedish
sott.wav	sutt.wav	1	N/A	naturaldur	reverse	Swedish
dytt.wav	t6tt.wav	1	N/A	naturaldur	normal	Swedish
gid.wav	ged.wav	1	nonce	naturaldur	normal	English
snudd.wav	snodd.wav	1	N/A	naturaldur	normal	Swedish
bed.wav	bad.wav	1	real	naturaldur	normal	English
glytt.wav	kludd.wav	1	N/A	naturaldur	normal	Swedish
mab.wav	mib.wav	1	nonce	naturaldur	reverse	English
msytt.wav	ms6tt.wav	1	N/A	manipdur	normal	Swedish

*Appendix E*

The stimulus list for the Trial Block 2 is as follows, with an ISI of 1000 ms.:

Stimulus 1	Stimulus2	ISI	Real/ Nonce	Manipulated?	Orientation	Language
mlydd.wav	ml6dd.wav	2	N/A	manipdur	normal	Swedish
tr6tt.wav	trytt.wav	2	N/A	naturaldur	reverse	Swedish
mutt.wav	mott.wav	2	N/A	naturaldur	normal	Swedish
mrytt.wav	mr6tt.wav	2	N/A	manipdur	normal	Swedish
sott.wav	sutt.wav	2	N/A	naturaldur	reverse	Swedish
trytt.wav	tr6tt.wav	2	N/A	naturaldur	normal	Swedish
led.wav	lad.wav	2	nonce	naturaldur	normal	English
mmib.wav	mmeb.wav	2	nonce	manipdur	normal	English
beg.wav	bag.wav	2	real	naturaldur	normal	English
dytt.wav	t6tt.wav	2	N/A	naturaldur	normal	Swedish
mgad.wav	mged.wav	2	nonce	manipdur	reverse	English
nytt.wav	n6tt.wav	2	N/A	naturaldur	normal	Swedish
mgytt.wav	mg6tt.wav	2	N/A	manipdur	normal	Swedish
ged.wav	gid.wav	2	nonce	naturaldur	reverse	English
gid.wav	ged.wav	2	nonce	naturaldur	normal	English
mgad.wav	mgid.wav	2	nonce	manipdur	reverse	English
mmott.wav	mmutt.wav	2	N/A	manipdur	reverse	Swedish
mmeb.wav	mmab.wav	2	nonce	manipdur	normal	English
big.wav	bag.wav	2	real	naturaldur	normal	English
meb.wav	mab.wav	2	nonce	naturaldur	normal	English
mbeg.wav	mbig.wav	2	real	manipdur	reverse	English
mf6tt.wav	mhytt.wav	2	N/A	manipdur	reverse	Swedish
mklytt.wav	mkl6tt.wav	2	N/A	manipdur	normal	Swedish
rott.wav	rutt.wav	2	N/A	naturaldur	reverse	Swedish

kludd.wav	glytt.wav	2	N/A	naturaldur	reverse	Swedish
mkeb.wav	mkib.wav	2	nonce	manipdur	reverse	English
mbed.wav	mbid.wav	2	real	manipdur	reverse	English
mfott.wav	mfutt.wav	2	N/A	manipdur	reverse	Swedish
lott.wav	lutt.wav	2	N/A	naturaldur	reverse	Swedish
bad.wav	bid.wav	2	real	naturaldur	reverse	English
msytt.wav	ms6tt.wav	2	N/A	manipdur	normal	Swedish
bid.wav	bed.wav	2	real	naturaldur	normal	English
mdytt.wav	mt6tt.wav	2	N/A	manipdur	normal	Swedish
mdib.wav	mdeb.wav	2	nonce	manipdur	normal	English
mtr6tt.wav	mtrytt.wav	2	N/A	manipdur	reverse	Swedish
mab.wav	mib.wav	2	nonce	naturaldur	reverse	English
mged.wav	mgad.wav	2	nonce	manipdur	normal	English
mdeb.wav	mdib.wav	2	nonce	manipdur	reverse	English
mrutt.wav	mdrott.wav	2	N/A	manipdur	normal	Swedish
gad.wav	gid.wav	2	nonce	naturaldur	reverse	English

*Appendix E*

The stimulus list for the Trial Block 3 is as follows, with an ISI of 500 ms.:

Stimulus 1	Stimulus 2	ISI	Real/ Nonce	Manipulated?	Orientation	Language
mrytt.wav	mr6tt.wav	1	N/A	manipdur	normal	Swedish
f6tt.wav	hytt.wav	1	N/A	naturaldur	reverse	Swedish
mbag.wav	mbig.wav	1	real	manipdur	reverse	English
kib.wav	kab.wav	1	nonce	naturaldur	normal	English
mmeb.wav	mmib.wav	1	nonce	manipdur	reverse	English
rytt.wav	r6tt.wav	1	N/A	naturaldur	normal	Swedish
tr6tt.wav	trytt.wav	1	N/A	naturaldur	reverse	Swedish
meb.wav	mab.wav	1	nonce	naturaldur	normal	English
mdrytt.wav	mdr6tt.wav	1	N/A	manipdur	normal	Swedish
bad.wav	bed.wav	1	real	naturaldur	reverse	English
led.wav	lid.wav	1	nonce	naturaldur	reverse	English
drott.wav	drutt.wav	1	N/A	naturaldur	reverse	Swedish
t6tt.wav	dytt.wav	1	N/A	naturaldur	reverse	Swedish
lad.wav	led.wav	1	nonce	naturaldur	reverse	English
mdr6tt.wav	mdrytt.wav	1	N/A	manipdur	reverse	Swedish
mled.wav	mlad.wav	1	nonce	manipdur	normal	English
mglytt.wav	mkludd.wav	1	N/A	manipdur	normal	Swedish
mbed.wav	mbid.wav	1	real	manipdur	reverse	English
mg6tt.wav	mgytt.wav	1	N/A	manipdur	reverse	Swedish
plutt.wav	glott.wav	1	N/A	naturaldur	normal	Swedish
nytt.wav	n6tt.wav	1	N/A	naturaldur	normal	Swedish
mgid.wav	mged.wav	1	nonce	manipdur	normal	English
mgad.wav	mged.wav	1	nonce	manipdur	reverse	English
mdib.wav	mdab.wav	1	nonce	manipdur	normal	English

ml6dd.wav	mlydd.wav	1	N/A	manipdur	reverse	Swedish
mgytt.wav	mg6tt.wav	1	N/A	manipdur	normal	Swedish
mjytt.wav	mj6tt.wav	1	N/A	manipdur	normal	Swedish
gad.wav	gid.wav	1	nonce	naturaldur	reverse	English
mdab.wav	mdib.wav	1	nonce	manipdur	reverse	English
lodd.wav	ludd.wav	1	N/A	naturaldur	reverse	Swedish
dab.wav	dib.wav	1	nonce	naturaldur	reverse	English
mbeg.wav	mbig.wav	1	real	manipdur	reverse	English
mbag.wav	mbeg.wav	1	real	manipdur	reverse	English
mbeg.wav	mbag.wav	1	real	manipdur	normal	English
kludd.wav	glytt.wav	1	N/A	naturaldur	reverse	Swedish
rutt.wav	rott.wav	1	N/A	naturaldur	normal	Swedish
jytt.wav	j6tt.wav	1	N/A	naturaldur	normal	Swedish
mr6tt.wav	mrytt.wav	1	N/A	manipdur	reverse	Swedish
trytt.wav	tr6tt.wav	1	N/A	naturaldur	normal	Swedish
klytt.wav	kl6tt.wav	1	N/A	naturaldur	normal	Swedish



*Appendix E*

The stimulus list for the Trial Block 4 is as follows, with an ISI of 1000 ms.:

Stimulus 1	Stimulus 2	ISI	Real/ Nonce	Manipulated?	Orientation	Language
futt.wav	fott.wav	2	N/A	naturaldur	normal	Swedish
glytt.wav	kludd.wav	2	N/A	naturaldur	normal	Swedish
trytt.wav	tr6tt.wav	2	N/A	naturaldur	normal	Swedish
mmutt.wav	mmott.wav	2	N/A	manipdur	normal	Swedish
knott.wav	knutt.wav	2	N/A	naturaldur	reverse	Swedish
mplutt.wav	mglott.wav	2	N/A	manipdur	normal	Swedish
mdytt.wav	mt6tt.wav	2	N/A	manipdur	normal	Swedish
mrutt.wav	mdrott.wav	2	N/A	manipdur	normal	Swedish
s6tt.wav	sytt.wav	2	N/A	naturaldur	reverse	Swedish
bed.wav	bad.wav	2	real	naturaldur	normal	English
mbid.wav	mbad.wav	2	real	manipdur	normal	English
mdrytt.wav	mdr6tt.wav	2	N/A	manipdur	normal	Swedish
mib.wav	mab.wav	2	nonce	naturaldur	normal	English
mdab.wav	mdeb.wav	2	nonce	manipdur	reverse	English
mutt.wav	mott.wav	2	N/A	naturaldur	normal	Swedish
dib.wav	dab.wav	2	nonce	naturaldur	normal	English
mknutt.wav	mknott.wav	2	N/A	manipdur	normal	Swedish
r6tt.wav	rytt.wav	2	N/A	naturaldur	reverse	Swedish
gid.wav	ged.wav	2	nonce	naturaldur	normal	English
mbag.wav	mbig.wav	2	real	manipdur	reverse	English
drytt.wav	dr6tt.wav	2	N/A	naturaldur	normal	Swedish
mbig.wav	mbeg.wav	2	real	manipdur	normal	English
mglott.wav	mplutt.wav	2	N/A	manipdur	reverse	Swedish
mgid.wav	mged.wav	2	nonce	manipdur	normal	English

mkib.wav	mkab.wav	2	nonce	manipdur	normal	English
sott.wav	sutt.wav	2	N/A	naturaldur	reverse	Swedish
mbeg.wav	mbig.wav	2	real	manipdur	reverse	English
mmab.wav	mmib.wav	2	nonce	manipdur	reverse	English
mlid.wav	mlad.wav	2	nonce	manipdur	normal	English
mbag.wav	mbeg.wav	2	real	manipdur	reverse	English
mgad.wav	mgid.wav	2	nonce	manipdur	reverse	English
mludd.wav	mlodd.wav	2	N/A	manipdur	normal	Swedish
t6tt.wav	dytt.wav	2	N/A	naturaldur	reverse	Swedish
mib.wav	meb.wav	2	nonce	naturaldur	normal	English
msutt.wav	msott.wav	2	N/A	manipdur	normal	Swedish
mlott.wav	mlutt.wav	2	N/A	manipdur	reverse	Swedish
rott.wav	rutt.wav	2	N/A	naturaldur	reverse	Swedish
dib.wav	deb.wav	2	nonce	naturaldur	normal	English
g6tt.wav	gytt.wav	2	N/A	naturaldur	reverse	Swedish
msk6tt.wav	mskytt.wav	2	N/A	manipdur	reverse	Swedish

*Appendix F*

Below are participants' decision results for study 2. Column headers are described below:

- Pair: Vowels found in a corresponding minimal pair
- Raw Score: Score obtained by coding participants' choices as either 0 (first word longer) or 1 (second word longer), and then totalling all decisions and calculating an average. Represents decisions in a percentage form.
- W1 longer: Number of times participants selected the first word in the pair as longer.
- W2 longer: Number of times participants selected the second word as longer.
- Total #: Total number of decisions made per each word pair.
- Chi-Square: Chi-Square value obtained from comparing W1 and W2 against the null hypothesis ( $Total\# / 2$ ). In all instances, degrees of freedom ( $df$ ) = 1.

**Chart 1a:** English Language; Natural Duration; Real Words; ISI = 500 ms.

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
i - e	0.6129	10	19	29	2.793	0.0947	
i - a	0.8709	4	25	29	15.207	<.0001	**
e - i	0.4516	17	14	31	0.29	0.59	
e - a	0.8709	4	25	29	15.207	<.0001	**
a - i	0.0967	26	3	29	18.241	<.0001	**
a - e	0.1129	55	7	62	37.161	<.0001	**

**Chart 1b:** English Language; Natural Duration; Real Words; ISI = 1000 ms.\*

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
i - e	0.6451	11	20	31	2.613	0.106	
i - a	0.9354	2	29	31	23.516	<.0001	**
e - a	0.8225	11	51	62	25.806	<.0001	**
a - i	0.1129	28	3	31	20.161	<.0001	**

\*Only pairs for which results were available have been included. As the number of stimuli was randomized and halved from the original list to shorten the trials, some results are not available.

*Appendix F*

**Chart 2a:** English Language; Natural Duration; Nonce Words; ISI = 500 ms.

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
i - e	0.7258	17	45	62	12.645	0.0004	**
i - a	0.8494	14	79	93	45.43	<.0001	**
e - i	0.4193	54	39	93	2.419	0.1198	
e - a	0.8924	10	83	93	57.301	<.0001	**
a - i	0.0967	84	9	93	60.484	<.0001	**
a - e	0.0322	30	1	31	27.129	<.0001	**

**Chart 2b:** English Language; Natural Duration; Nonce Words; ISI = 1000 ms.\*

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
i - e	0.6693	41	83	124	14.226	0.0002	**
i - a	0.8870	7	55	62	37.161	<.0001	**
e - i	0.1612	26	5	31	14.226	0.0002	**
e - a	0.9516	3	59	62	50.851	<.0001	**
a - i	0.1290	54	8	62	34.129	<.0001	**

\*Only pairs for which results were available have been included.

**Chart 3a:** English Language; Manipulated Duration; Real Words; ISI = 500 ms.\*

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
i - e	0.8064	4	25	29	15.207	<.0001	**
e - i	0.4838	40	45	85	0.294	0.5876	
e - a	0.8709	4	27	31	17.065	<.0001	**
a - i	0.2419	47	15	62	16.516	<.0001	**
a - e	0.1935	25	6	31	11.645	0.0006	**

\*Only pairs for which results were available have been included.

*Appendix F*

**Chart 3b:** English Language; Manipulated Duration; Real Words; ISI = 1000 ms.\*

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
i - e	0.5483	14	17	31	0.29	0.59	
i - a	0.9354	1	29	30	26.133	<.0001	**
e - i	0.4193	54	39	93	2.419	0.1198	
a - i	0.2580	23	8	31	7.258	0.0071	**
a - e	0.1290	27	4	31	17.065	<.0001	**

\*Only pairs for which results were available have been included.

**Chart 4a:** English Language; Manipulated Duration; Nonce Words; ISI = 500 ms.

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
i - e	0.7258	17	45	62	1.645	0.0004	**
i - a	0.8924	10	83	93	57.31	<.0001	**
e - i	0.3548	80	44	124	10.452	0.0012	**
e - a	0.8870	7	55	62	37.161	<.0001	**
a - i	0.1129	55	7	62	37.161	<.0001	**
a - e	0.1612	52	10	62	28.452	<.0001	**

**Chart 4b:** English Language; Manipulated Duration; Nonce Words; ISI = 1000 ms.

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
i - e	0.6021	37	56	93	3.882	0.0488	*
i - a	0.8709	8	54	62	34.129	<.0001	**
e - i	0.3709	39	23	62	4.129	0.0422	*
e - a	0.9354	4	58	62	47.032	<.0001	**

a - i	0.1397	80	13	93	48.269	<.0001	**
a - e	0.1129	55	7	62	37.161	<.0001	**

*Appendix F*

**Chart 5a:** Swedish Language; Natural Duration; ISI = 500 ms.

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
y - ø	0.7060	82	197	279	47.401	<.0001	**
ø - y	0.3494	121	65	186	16.86	<.0001	**
u - o	0.6774	40	84	124	15.613	<.0001	**
o - u	0.3870	76	48	124	6.323	0.0119	*

**Chart 5b:** Swedish Language; Natural Duration; ISI = 1000 ms.

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
y - ø	0.7612	37	118	155	42.329	<.0001	**
ø - y	0.2774	112	43	155	30.716	<.0001	**
u - o	0.8279	16	77	93	40.011	<.0001	**
o - u	0.4516	102	84	186	1.742	0.1869	

**Chart 6a:** Swedish Language; Manipulated Duration; ISI = 500 ms.

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
y - ø	0.7871	33	122	155	51.103	<.0001	**
ø - y	0.4014	167	112	279	10.842	0.001	**
o - u	0.3483	101	54	155	14.252	0.0002	**

\*Only pairs for which results were available have been included.

**Chart 6b:** Swedish Language; Manipulated Duration; ISI = 1000 ms.

Pair	Raw Score	W1 longer	W2 longer	Total #	Chi-Square	p-value	sig. ( $\alpha=.01$ )
y - ø	0.7701	57	191	248	72.403	<.0001	**

ø - y	0.4946	47	46	93	0.011	0.9174	
u - o	0.7788	48	169	217	67.47	<.0001	**
o - u	0.3790	77	47	124	7.258	0.0071	**

*Appendix G*

*Phonological neighbourhood densities.* These measures are important for understanding how closely linked nonce words are to similar-sounding real words.

Nonce Word	# Neighbours
/dɪb/	30
/dɛb/	20
/dæb/	27
/mɪb/	23
/mɛb/	17
/mæb/	26
/gɪd/	28
/gɛd/	25
/gæd/	37
/kɪb/	23
/kɛb/	14
/kæb/	30

## *Appendix H*



### **Invitación de participar en un estudio de la fonética**

Título de proyecto de investigación: La Altura y Duración de Vocales inglesas y españolas

Fecha de inicio de la autorización del ética: El 8 de Agosto de 2014

La autorización del ética para la colección de datos expirará: El 31 de Mayo de 2015

Estimado Señor / Señora/ Señorita,

Somos investigadores del Instituto de ciencia cognitiva buscando voluntarios de la Universidad de Carleton. Mi nombre es Adam Stone (Adam\_stone@carleton.ca) y actualmente soy el candidato de estudiante de posgrado, trabajando bajo la supervisión del Dr. Ida Toivonen (Ida\_Toivonen@carleton.ca). Necesitamos voluntarios para participar en un estudio que examina la calidad y cantidad de vocales españolas chilenas. Se le pedirá a leer una lista de palabras, y su voz se grabará. Cada palabra se leerá en una frase. El estudio no llevará más de veinte minutos. Usted no será compensado en cualquier forma.

Su información se mantendrá estrictamente confidencial y anónimo. Nadie, aparte de los investigadores, tendrá acceso a tu información. Los investigadores no utilizará ninguna información de identificación cuando se realizan análisis de datos. Usted no será identificado con sus resultados. Los investigadores sólo utilizará sus datos para los propósitos de este estudio

Debe usted decidió participar en este estudio, su participación puede ser retirada en cualquier momento hasta la terminación de la grabación oral. Por favor tenga en cuenta que esto es un estudio de riesgo



mínimo y no pretende causar ningún daño a usted como participante.

Este proyecto ha sido revisado y aprobado por el Consejo de ética de investigación de la Universidad Carleton. Si tienes alguna duda o pregunta acerca de este estudio, por favor póngase en contacto con los presidentes del Consejo de ética de investigación en:

*Appendix H*

Professor Andy Adler, Chair

Research Ethics Board

Carleton University Research Office

Carleton University

1125 Colonel By Drive

Ottawa, Ontario K1S 5B6

Tel: 613-520-2517 E-mail: [ethics@carleton.ca](mailto:ethics@carleton.ca)

Si están dispuestos a participar en este estudio y son por lo menos 18 años de edad o más, por favor póngase en contacto con Adam Stone en [adam\\_stone@carleton.ca](mailto:adam_stone@carleton.ca)

Sinceramente,

Adam Stone

*Appendix H*



Consentimiento del participante

Título: La altura y la duración de vocales en inglés y español

Fecha de inicio de la autorización: El 8 de Agosto de 2014

Autorización del ética para la colección de datos expira: El 31 de Mayo de 2015

Yo \_\_\_\_\_ deciden participar en un estudio sobre la altura de la vocal. Este estudio pretende investigar la calidad y cantidad de las vocales en español chileno. Mi participación constará de un número de palabras leído en voz alta. Las palabras se pronunciarán en una frase del portador, así el investigador dice "decirme X", donde X es la palabra objetivo. Este proyecto se lleva a cabo por Adam Stone bajo la supervisión del Dr. Ida Toivonen.

El participante tiene derecho a poner fin a su participación en el estudio en cualquier momento, hasta la terminación del ejercicio oral grabado. Debido a la promesa del anonimato, no es posible retirar después de terminar la grabación. Los participantes que retiran del estudio tendrán su información inmediatamente destruido.

Los participantes permanecerán en el anonimato en este estudio. Los nombres y otros identificadores personales no se atribuyen a las observaciones. El investigador del supervisor, el Dr. Ida Toivonen, pueden tener acceso a las notas de campo, pero no más allá del investigador y su supervisor tendrán acceso. Confidencialidad se destacó en todo momento.

Las notas se archivarán después de la terminación del proyecto de investigación. Esto significa que las notas se mantendrá en un gabinete cerrado si realizarse unas investigaciones adicionales por el

investigador sobre este tema en el futuro. Los datos no se utilizará para ningún propósito más allá de un estudio de la altura de la vocal.

### *Appendix H*

Este proyecto fue examinado por el Consejo del Ética de Investigación del Universidad Carleton, que proporcionó la autorización para realizar la investigación. Si tiene preguntas o preocupaciones a su participación en esta investigación, por favor póngase en contacto:

Profesor Andy Adler, Chair  
Research Ethics Board  
Carleton University  
1325 Dunton Tower  
1125 Colonel By Drive  
Ottawa, ON K1S 5B6  
Tel: 613-520-2517

**Researcher contact information:**

Adam William Stone  
Institute of Cognitive Science  
Carleton University  
Tel: 613-816-9112  
Email: adam\_stone@carleton.ca

**Supervisor contact information:**

Dr. Ida Toivonen  
Institute of Cognitive Science  
Carleton University  
Tel: 613-520-2600  
Email: Ida\_Toivonen@carleton.ca

Mediante la firma de éste, se entenderá que usted haya consentido a participar en el proyecto, y que usted da su consentimiento a la publicación de los resultados con el entendimiento de que se mantendrá el anonimato. Si usted acepta participar en este estudio y es mayores de 18 años, por favor firme a continuación.

---

---

Firma del participante

Fecha

---

Firma del investigador

---

Fecha

*Appendix I*



**Invitation to participate in a phonetics perception study**

Title of research project: The Height and Duration of English and Spanish Vowels

Date of ethics clearance: August 8, 2014

Ethics Clearance for the Collection of Data Expires: May 31, 2015

Dear Sir/Madam,

We are researchers from the Institute of Cognitive Science at Carleton University seeking volunteers. My name is Adam Stone ([adam\\_stone@carleton.ca](mailto:adam_stone@carleton.ca)) and I am currently a Master's Degree candidate, working under the supervision of Dr. Ida Toivonen ([Ida\\_Toivonen@carleton.ca](mailto:Ida_Toivonen@carleton.ca)). We need volunteers to participate in a study that examines the quality and quantity of English vowels. You will be asked to listen to word pairs through headphones on a computer, and choose which word out of the pair is the longest. The study will not take more than twenty minutes. You will be compensated in the form of \$5 Tim Horton's Gift Certificates.

Your information will be kept strictly confidential and anonymous. No one, other than the researchers, will have access to your information. The researchers will not use any identifying information when conducting data analysis. You will not be identified with your results. The researchers will only use your data for the purposes of this study.

Should you decide to participate in this study, your participation can be withdrawn at any time to the completion of the oral recording. Please note that this is a minimal risk study and is not expected to cause any harm to you as a participant.

This project has been reviewed and approved by the Carleton University Research Ethics Board. Should you have any concerns or questions about this study, please contact the Research Ethics Board chairs at:

*Appendix I*

Professor Andy Adler, Chair  
Research Ethics Board  
Carleton University Research Office  
Carleton University  
1125 Colonel By Drive  
Ottawa, Ontario K1S 5B6  
Tel: 613-520-2517 E-mail: [ethics@carleton.ca](mailto:ethics@carleton.ca)

If you are willing to partake in this study, and are at least 18 years of age or older, please contact Adam Stone at [adam\\_stone@carleton.ca](mailto:adam_stone@carleton.ca)

Sincerely,

Adam Stone

*Appendix I*



**Consent of the participant**

Title: The height and duration of English and Spanish vowels

Date of ethics clearance: August 08, 2014

Ethics Clearance for the Collection of Data Expires: May 31, 2015

I \_\_\_\_\_, choose to participate in a study on vowel height. This study aims to investigate the quality and quantity of vowels. My participation will consist of hearing words spoken in headphones, and making choices on a computer screen by pressing color-coded buttons. This project is carried out by Adam Stone under the supervision of Dr. Ida Toivonen.

---

The participant has the right to end his or her participation in the study at any time, up to the completion of the recorded oral exercise. Due to the promise of anonymity, it is not possible to withdraw after the recording is complete. Participants who withdraw from the study will have their information immediately destroyed.

Participants will remain anonymous in this study. Names and other personal identifiers will not be attributed to the observations noted. The researcher's supervisor, Dr. Ida Toivonen, may have access to the field notes but no one beyond the researcher and his supervisor will have access. Confidentiality will be stressed at all times.

The notes will be archived after completion of the research project. This means that the notes will be kept in a locked cabinet should any further research be done by the researcher on this topic. The data will not be used for any purpose beyond a study of vowel height.

This project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research. Should you have questions or concerns related to your involvement in this research, please contact:

*Appendix I*

Professor Andy Adler, Chair  
Research Ethics Board  
Carleton University  
1325 Dunton Tower  
1125 Colonel By Drive  
Ottawa, ON K1S 5B6  
Tel: 613-520-2517

**Researcher contact information:**

Adam William Stone  
Institute of Cognitive Science  
Carleton University  
Tel: 613-816-9112  
Email: adam\_stone@carleton.ca

**Supervisor contact information:**

Dr. Ida Toivonen  
Institute of Cognitive Science  
Carleton University  
Tel: 613-520-2600  
Email: Ida\_Toivonen@carleton.ca

By signing this, it will be understood that you have consented to participate in the project, and that you consent to publication of the results with the understanding that anonymity will be preserved. If you agree to participate in this study and are over the age of 18, please sign below.

\_\_\_\_\_  
Signature of participant

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature of researcher

\_\_\_\_\_  
Date

