Examining L2 vowel contrast desensitization in the EFL Classroom: A study on native Spanish/Catalan speaker perception and production of English tense-lax vowel pairs

By

Adrián Salcedo Silva

Submitted to the

Department of English and German Studies

in partial fulfillment of the requirements for the degree of

Master's in Teaching and Learning English as a Foreign / Second Language

at

UNIVERSITAT ROVIRA I VIRGILI

June 13, 2024

Signature of Student:

Adrián Salcedo Silva

Digitally signed by Adrián Salcedo Silva
Date: 2024.06.13 23:42:27 +02'00'

....

Certified by Dr. /Ms./Mr. ...., Master's Thesis Supervisor

Signature of Supervisor:

Joaquín Romero

Digitally signed by Joaquín Romero
Date: 2024.06.13 23:39:31 +02'00'

....

UNIVERSITAT ROVIRA I VIRGILI

MATLEFS

Master's in Teaching and Learning English as a Foreign/Second Language

2023-2024

Master's Final Project
ORIGINAL WORK FORM

STUDENT

Last Name(s) Salcedo Silva

First Name Adrián

I hereby state that all the work presented as part of my Final Master's Project is original with no exception and that I have not, voluntarily or otherwise, misused or misreported any previously published information. I am aware that any failures to comply with these statements will automatically result in disqualification of my final paper and that I will not be able to obtain any credits for it.

Signature of Student

In Tarragona, June 12 2024.

PLACE MONTH DAY

Abstract

This study aimed to assess the cue-weighting of duration and spectra in non-native vowel perception and production by Spanish/Catalan EFL learners of 13-14 years of age. A synthetic /i/-/ɪ/ continuum was created with five equal spectral and duration steps, that were repeated and randomised in a forced-choice identification task. The stimuli were collected from a native speaker and manipulated. The sample population (N = 43) was divided into three groups according to pronunciation training. Each group received four hours of pronunciation training based on articulatory activities, vowel-discrimination activities, and acoustic descriptions. One group was taught explicitly (N= 10), based on tenseness categorisations, another group learnt pronunciation implicitly (N = 15) and another learnt length-based vowel categorisations (N = 18). The perception data were analysed using Linear Mixed Models to find main effects for duration and quality manipulations, as well as estimated marginal means. Production data was also analysed using Linear Mixed Models and compared with spectral and duration values of natives as a measure of native-likeness. Production data showed partial reliance on duration at some steps of the continuum, and to a lesser extent, on quality contrasts. Production data found incipient contrastive categorisation of non-native tense-lax phonemes /i/-/ɪ/ and /u/-/ʊ/.

Keywords: *desensitization hypothesis, cue-weighting of duration, phonological acquisition, tense, lax*

Acknowledgements

I am deeply indebted to all those who have helped and supported me throughout the process of this research.

Firstly, I am immensely grateful to my supervisor, Joaquín Romero, for his continuing guidance and support, insightful feedback, and constant encouragement. His expertise and patience were invaluable in shaping this study.

Furthermore, I want to extend my deepest gratitude to all those participants from my internship who volunteered to take part in my research. I am also truly grateful to my classmates and especially my American classmates and Marni Manegre who helped me with the experiment's stimuli.

Last but, not least I want to thank all my family and my girlfriend, Diana Vasile, for their unwavering support and wise words.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

This dissertation aims to ascertain the cue-weighting of duration in vowel perception among non-native Spanish/Catalan speakers of English. Central to this thesis is Bohn's (1995) desensitization hypothesis (DH) which "states that whenever spectral differences are insufficient to differentiate vowel contrasts because previous linguistic experience did not sensitize listeners to these spectral differences, duration differences will be used to differentiate the non-native vowel contrast" (p. 294).

In effect, Bohn's (1995) hypothesis ascribes the speaker's reliance on duration contrasts to the lack of similar phonological categories in the speaker's vowel inventory and vowel space. For example, while the vowel inventory of an English speaker has two vowel categories in the high front vowel space (tense and lax /i/-/ɪ/), that region of the vowel space of a Spanish speaker only contains one category (viz. /i/), thus desensitizing Spanish speakers to phonological spectral differences in the high front region.

This thesis, therefore, aims to ascertain the extent to which Bohn's hypothesis accounts for the non-native perception of English tense lax contrast that Spanish and Catalan lack. This thesis also aims to find and compare the effectiveness of three approaches to pronunciation instruction: length-based, quality-based, and implicit.

The empirical support for Bohn's hypothesis has largely neglected the population of non-native speakers who have not received significant pronunciation training and who have not spent considerable time in an English-speaking country (being exposed to L2 input).

1.1 Background

1.1.1 Phonetic Cue weighting in Production and Perception

In speech perception, a phonetic cue may broadly refer to "any information that systematically influences listener's perception of a contrast" (Schertz & Clare, 2020, p. 2). However, Wright (2004, as cited in Schertz & Clare, 2020, p. 2) provides a narrower definition of cue weighting, "information in the acoustic signal that allows the listener to apprehend the existence of a phonological contrast." The fundamental difference between both definitions is that the narrower interpretation of phonetic cue considers phonological contrasts an essential characteristic of phonetic cues, and so most studies on cue weighting adopt this vision.

In this context, cue weighting refers to the significance of a cue to phonological contrasts either in production or in perception. These phonetic cues can be essential to distinguishing between two different lexical items by means of specific acoustic characteristics unconsciously acquired by speakers. For example, the voice onset time (VOT) can be the decisive cue that helps distinguish aspirated from unaspirated plosives in items like *beat-Pete*, where the word-initial plosives differ in VOT (Schertz & Clare, 2020). While VOT can be considered a "primary cue", because it is the most reliable cue for distinguishing phonemes, "secondary cues" can also be present in phonemic contrasts, although they will not carry as much phonological significance as primary cues. For example, /b/ and /p/ primarily differ in VOT but can also differ in F0 (fundamental frequency); however, F0 alone cannot distinguish between /b/ and /p/, whereas VOT can.

1.1.2 Tense-lax contrasts in English.

In the English language, there exists a vowel contrast referred to as "tenseness". This aspect of English vowels can be understood in articulatory terms as tongue-body tension caused by additional muscular effort (Wayland, 2018). For example, English high front vowels differ in tenseness /i/ and /ɪ/, that is, /i/ is said to be tense because in pronouncing it, there is more muscular effort involved than in pronouncing lax /ɪ/. Other vowels that differ in tenseness are the high back tense and lax vowels /u/-/ʊ/. However, these vowel pairs also vary in quality and duration. The tense counterparts are higher and tend to be longer than their lax counterparts in the same contexts. Nevertheless, the primary cue in distinguishing tense and lax vowels is quality (spectra), not quantity because lax vowels can be longer than their tense counterparts; however, tense vowels will always be spectrally different than their lax counterparts. Despite the primary weighting of spectra for tense and lax vowels discrimination, length-based classifications of tense and lax vowels have partly characterised British approaches to English phonology, whereas American approaches have tended to emphasise the role of tenseness (tension) and their distributional correlate, in distinguishing tense and lax vowels. Length-based classifications, however, may be misleading in that "long" vowels may actually be shorter and "short" vowels, especially when preceded by word-final voiceless plosives.

1.2 Theoretical Background

Seminal studies on second language (L2) speech perception have consistently documented a key finding in L2 research: the influence of speakers' native language (L1) on their perception of L2 speech and categorisation of non-native sounds. Perhaps this finding may be best summarised as the transfer hypothesis which Trubetzkoy (1939) already envisioned in stating that "[t]he phonological system of a language [L1] is like a sieve through which everything that is said passes" (p. 51). Later studies also attested that a great many non-native perception and production errors can be attributed to L1 transfer (e.g., Bent & Bradlow, 2003; Chung & Kim, 2021; Flege & Hillenbrand, 1984; Morrison, 2003).

This section will describe two prominent L2 speech perception models, namely, the Revised Speech Learning Model (SML-r, Flege & Bohn, 2021) and the Perceptual Assimilation Model as applied to L2 speech perception (PAM-L2, Best & Tyler, 2007). Additionally, this section will also provide a detailed account of Bohn's (1995) desensitization hypothesis (DH), which contrasts with the L2 models mentioned above in that it offers an L1-independent account of L2 vowel perception.

The SLM-r establishes the main factors that determine the success of acquiring L2 vowel categories. One of these factors is the quantity and quality of L2 input to which learners are exposed; the second factor is the perception of the closeness between the L2 sound and the most similar L1 sound; and the third factor is the configuration of the L1 categories at the onset of L2 learning.

The first difference between the SLM and SLM-r is that the revised model does not consider the contrasts between young and old learners for speech learning. Secondly, the revised version assumes the unlikely native-like perception and production of L2 sounds among non-native learners, forgoing the theoretical relevance of highly proficient L2 individuals (Flege & Bohn, 2021).

The second model (PAM-L2) supports an assimilation pattern of L2 sounds. This means that learners' perceptions of L2 sounds are invariably affected by their L1's phonological categories, making similarities and dissimilarities (between L1 and L2 categories) determining factors in the accurate perception of L2 sounds.

In effect, these models underscore the importance of L2 input and environment in the acquisition of the L2. They also focus on the importance of the similarity and closeness of the L1 and L2 phonological categories, which will predict the accurate discriminability of L2 categories to a certain extent.

1.3 The Desensitization Hypothesis (DH)

Bohn (1995) formulated the desensitization hypothesis (DH) to account for the phenomenon observed in a number of vowel perception studies (Bennett 1968; Weiss 1976b; Lieberman & Kubaska, 1979; Butcher, 1976; Flege & Bohn, 1989; Bohn and Flege, 1990; Flege & Bohn, 1994, as cited in Bohn, 1995). These studies indicated that listeners relied on duration as the main phonetic cue in contrasting non-native vowel sounds, irrespective of the listeners' native language (L1) experience with duration as a significant phonetic cue.

Specifically, the desensitization hypothesis (DH) posits that "whenever spectral differences are insufficient to differentiate vowel contrasts because previous linguistic experience did not sensitize listeners to these spectral differences, duration differences will be used to differentiate the non-native vowel contrast." (Bohn, 1995, p. 294). Later, Bohn (2019)reformulated this postulation by appending "irrespective of whether the duration cue is phonologically relevant in the listener's L1" (p. 5). Thus, this hypothesis is predicated on the assumption that L1 background (native language sound system) may not always influence non-native vowel perception, unlike what traditional L1 transfer accounts may claim. Traditional L1 transfer hypotheses suggest that listeners rely on durational cues in non-native vowel contrasts because they use these strategies in their L1. In contrast, the DH attributes this reliance to the *absence* of sensitization to specific spectral cues in the vowel space. Crucially, the DH explains why some listeners rely on duration even if they do not have experience with this cue in contrasting vowels in their L1.

In essence, in the absence of linguistic experience with spectral contrasts in a specific region of the vowel space (e.g., in English the high front region has tense and lax contrasting vowels /i/ and /ɪ/), durational contrasts will take precedence over spectral contrasts even though the latter may be the most familiar and reliable cue in the listener's L1. In this context, the term "desensitization" refers to the loss of sensitivity to spectral cues (quality) in non-native vowel contrasts.

Importantly, the DH sees this reliance on duration as a manifestation of a universal perceptual tendency irrespective of L1, observed across multiple perception studies. Listeners of different sound systems, Spanish and Mandarin (Flege & Bohn, 1989; Flege, Wang, & Bohn, n.d., as cited in Bohn, 1995), were observed to rely on duration in perceiving non-native vowels that were in proximity in the vowel space (i.e., /i/ and /ɪ/). Even though the vowel arrangement of Spanish and Mandrin is different and both systems rely fundamentally on spectral contrasts to differentiate vowels, both groups of listeners seemed to find durational cues easier to access when identifying the vowels /i/ and /ɪ/.

In his opening remarks on vowel perception studies, Bohn (1995) proposes a similar change in focus as that seen in production studies. He advocates for perception studies to move their attention from subject variables, such as L1 background, to target variables such as "acoustic cues" and "relation of non-native contrasts to native categories" (Bohn, 1995, p. 280). Bohn argues this change is necessary because cross-language perception studies that rely on subject variables fail to satisfactorily explain phenomena related to non-native phonetic cue-weighting of duration, independent of L1 background.

Importantly, Bohn does not challenge the influence of listeners' L1, but merely suggests that it may not be the sole conditioning factor in non-native perception. In this regard, Bohn proposes the DH as complementary rather than contradictory to previous cross-language theories based on L1 influence and transfer.

An early study on cue weighting in vowel perception that introduced the target variables mentioned above was conducted by Gottfried and Beddor (1988). Their experiment involved three groups of listeners: native French, native American English who were French learners, and native American English who were not French learners. These groups were administered an identification test consisting of a synthetic continuum of the French vowels /o/-/ɔ/ in the context of /kVt/ (using the words c*ôte* and *cotte*, respectively). The synthetic continuum was created manipulating the F1 and F2 values by an increment of 5 Hz from c*ôte (*/kot/) to *cotte* (/kɔt/), resulting in 10 steps of spectrally differing stimuli. For each of the spectrally manipulated stimuli, three other temporal (duration) variations were created by increasing the duration of the vowel nuclei by 40 (ms): 140 ms, 180 ms, and 240 ms. The manipulations resulted in 30 stimuli in total.

The study revealed that the native English groups predominantly relied on vowel duration in their phonemic categorisation of the French vowels /o/-/ɔ/. In contrast, the native French group were shown to attend to spectral contrasts more than temporal ones.

In effect, the native English group often identified long vowels as /o/ and short vowels as /ɔ/ when they were spectrally ambiguous (mid-point of the continuum).

At the time, Gottfried and Beddor (1988) interpreted these findings through the lens of L1 interference (i.e., L1 background). They argued that the non-native group's reliance on duration was a "function of the systematic role of duration in differentiating spectrally adjacent vowels in English." (1988, p. 62). Thus, the authors attributed the native English group's reliance on duration to their experience with this cue in their L1, despite the higher significance of spectrum as a contrasting cue in English (cf. Bohn & Flege, 1990; Flege et al., 1997; Hillenbrand et al., 2000; Wayland, 2018; Zhu et al., 2023).

Conversely, Gottfried and Beddor (1988) surmised that the French listeners' predominant reliance on spectra was due to the limited experience of duration as a significant cue. The /o/-/ɔ/ is the only vowel contrast that exhibits covariance between spectra and duration in French.

Although Gottfried and Beddor's experiment may not have yielded conclusive evidence for the DH, given the purported influence of duration in English, later studies by Bohn and Flege demonstrated how L1 transfer hypotheses are insufficient to account for *all* instances of cross-language vowel perception (Flege & Bohn, 1989; Bohn & Flege 1990; Flege & Bohn 1994, as cited in Bohn, 1995). In particular, Bohn and Flege conducted three experiments with different groups of native speakers of German, Spanish, and Mandarin that resembled that of Gottfried and Beddor in its forced-choice identification design (the experimental design is reported in Bohn & Flege, 1990).

In the three experiments, synthetic continua were created that contained 11 spectrally different stimuli: from /i/ to /ɪ/ (using the words "beat" and "bit", respectively) for the Spanish and Mandarin groups, and from /ɛ/ to /æ/ (using the words "bet" and "bat", respectively) for the German group. For each of the stimuli of the continua, durations were also manipulated in increments of 50 ms (150 ms, 200 ms, and 250 ms). The experiments aimed to assess the cue weighting of spectrum and duration of non-native vowel perception in Spanish, Mandarin, and German listeners.

The rationale behind the choice of vowel pairs and languages is that Spanish, Mandarin, and German listeners are sensitised to a region of their vowel space where they have only one vowel category, whereas English has two. This configuration enabled the researchers to ascertain the role of spectra and duration in categorising and contrasting non-native vowels adjacent to native vowels in the vowel space (Bohn, 1995).

Bohn (1995, p. 286) describes illustrations of vowel spaces that depict how Spanish and Mandarin share roughly the same and only vowel category in the high front region of the acoustic vowel space (i.e., /i/), while English has two different categories in that region, contrasting in tenseness: namely, tense /i/ and lax /ɪ/. Likewise, the second illustration displays the approximate placement of English /ɛ/ and /æ/ relative to the German vowels /ɛ/ and /a/. The German /ɛ/ overlaps with its English counterpart /ɛ/, whereas the English /æ/ has no clear correspondence in German.

Bohn (1995) interpreted the results of the first experiment involving German listeners as supporting L1 transfer hypotheses and their predictions. Given the greater significance of duration as a phonetic cue in German than in English, it would make sense that the German listeners transferred their native perceptual strategies with duration to their non-native vowel perception maintaining the same complementarity with spectra as found in their native system.

Bohn (1995, p. 289) reported the results of the identification scores for the continuum /ɛ/ - /æ/ from the English and German participants by means of graphs that plotted "% Identification as 'bet'" as the y-axis and as a function of spectral steps (x-axis) and duration steps (different lines). Bohn used this illustration to reveal how English participants relied mainly on spectra, given the almost constant overlap of identification functions for length variations (short, medium, and long vowels). However, the German listeners relied on duration, as inferred from the stark contrast in identification functions for duration variations. The identifications gaps between long, medium, and short stimuli are maintained throughout the continuum.

On the other hand, the results of the experiments with Spanish and Mandarin participants did not meet the predictions of L1 transfer hypotheses (i.e., relying mainly on spectra, as this is the main cue in Spanish and Mandarin). Instead, Spanish participants predominantly relied on duration, and Mandarin listeners almost relied on this cue alone to identify tense /i/ and lax /ɪ/.

The results of the three experiments revealed the need for an alternative account to those that attributed non-native vowel perception solely to subject variables like L1 background. While the German group's reliance on duration could be ascribed to their experience with this cue in their L1, the Spanish and Mandarin groups' predominant reliance on duration cannot be solely explained by L1 transfer, since duration is non-phonemic at a segmental level in Spanish or Mandarin.[1] The desensitisation hypothesis, however, affords an interpretation of these results by pointing to a language-independent

[1] Although duration is not significant in Mandarin vowel contrasts, it has been shown to consistently pattern with two of the four tones of Mandarin (Nordenhake & Svantesson, 1983, as cited in Bohn, 1995). Bohn, then, considers that experience with duration as a concomitant feature on a suprasegmental level may have sensitised listeners to duration contrasts in non-native vowel perception.

preference for duration as the main contrasting cue when listeners have not been provided with sufficient experience with specific spectral contrasts.

Although Bohn (1995) de-emphasises the role of L1 background as a determiner of cross-language vowel perception in his desensitization hypothesis, he claims that L1 experience has an important role in (de)sensitising listeners to specific areas of their acoustic vowel space. In this regard, Bohn proposes that spectral (de)sensitisation may operate as a function of L1 experience. In support of this argument, Bohn refers to Butcher's (1976, as cited in Bohn, 1995) study on cross-language cardinal vowel perception with British and German listeners of different age groups.

Part of Butcher's (1976, as cited in Bohn, 1995) experiment consisted of a vowel dissimilarity rating task with front vowels /ɛ/ and /a/. One of the main findings was that the adult British group perceived greater distances between the two vowels than did the adult German group. Additionally, when compared with other age groups, the British children perceived smaller differences between the vowels than did the adult British group; conversely, the German children perceived greater differences than the adult German group.

Butcher (1976, as cited in Bohn, 1995) attributed the findings from the adult groups to the fact that vowel categories in the low front region are closer for the English vowel space than for the German vowel space. Thus, the closeness that English vowels exhibit may have enabled British speakers to discriminate vowels more accurately.

Crucially, the findings from the comparison between age groups were interpreted by Bohn as manifesting a "(de)sensitisation effect" attributed to the arrangement of native vowel categories and L1 experience. Bohn argued that, as English speakers become more linguistically experienced, they are sensitised to spectral contrasts in the low front region, since their vowel space exhibits closer vowel categories in that region, and thus greater discriminability. On the other hand, linguistic experience desensitises German speakers to spectral contrasts in the low front region because their native vowel categories are further apart in that area, priming them to resort to duration, not spectrum, to distinguish vowel contrasts.

In light of the findings from Butcher (1976, as cited in Bohn, 1995) and the postulation of the desensitization hypothesis, Bohn (1995) provides a revised interpretation of the results from Gottfried and Beddor (1988) with French and English participants, as well as Bohn and Flege's (1990) experiment with German and English participants. In both studies, Bohn attributes the non-native participants' reliance on duration to the lack of experience with the spectral contrasts of such foreign vowels.

For example, in Gottfried and Beddor (1988), American English speakers relied on duration "because most speakers of American English lack an /ɔ/ category" (Bohn, 1995, p. 297), thus desensitising American speakers to spectral contrasts in the back region of the vowel space. Similarly, in Bohn and Flege (1990), German speakers' reliance on duration could be accounted for by the desensitization hypothesis, as Germans lack a /ɛ/ - /æ/ contrast and "never needed to direct their attention to spectral differences in that area [low front]" (Bohn, 1995, p. 295).

1.3.1 An alternative account to the desensitization hypothesis

Despite the empirical evidence for the desensitization hypothesis, Bohn (1995) identifies an alternate account for the significant reliance on duration observed in numerous studies. Bohn suggests that participants in previous studies, and those described above, may have been primed to attend to duration differences because of the design of the experiments. In particular, Bohn surmises that the unequal number of manipulated spectra and duration steps may have prompted participants to perceive duration more readily than spectra, since the former varied the least.

In the context of the experiments described above, the spectral dimension of the continua in Bohn (1995) resulted in 11 stimuli, whereas duration manipulations resulted in 3 steps. Likewise, in Gottfried and Beddor's (1988) experiment, the stimuli exhibited unequal number of manipulations, with 10 spectrally manipulated stimuli and only 3 duration manipulations.

1.4 Acquisition of English phonology by English/Catalan speakers

The relevant literature has extensively dealt with the perception and acquisition of L2 vowel sounds (including in EFL settings with native Catalan speakers). However, Aliaga-García's (2017) research is especially relevant to the aim of this thesis, since it also aims to assess the effectiveness of training methods for improving the perception of non-native vowel pairs. However, her study differs in that the "desensitization hypothesis" takes a more significant role in this research, while Aliaga-García's research examines the hypothesis from a broader perspective, and this research includes young participants from secondary school, unlike her research, which includes adult learners.

Another similar study is Cerviño-Povedano and Mora's (2015) research on how Spanish EFL learners categorize the tense-lax vowel pair /i/-/ɪ/. While the study also focuses on the cue-weighting of spectral and durational aspects in the perception of these vowels, the study does not consider the desensitization hypothesis; instead, it addresses the perception of these vowels in terms of phonological short-term memory (PSTM). This study is significant because it shows how Spanish learners rely mainly on durational patterns instead of spectral ones when discriminating vowel pairs. The present study will consider the desensitization hypothesis as a potential description of this phenomenon.

This thesis also shares similarities with Souza's et al. (2017) work in that they also take into account Bohn's hypothesis in their analysis. However, the differences lie in the scope of the languages compared. While this study focuses on Catalan/Spanish speakers, Souza et al. examine Russian, Danish, Catalan, and Portuguese in their perception of the tense-lax high front vowel pair. Moreover, contrary to this study, they do not assess the effectiveness of training methods for improving L2 vowel perceptions. Instead, they aim to relate vowel inventory size with reliance on durational cues, as opposed to spectral ones.

The most recent similar study is Yu's (2023) research on the extent to which duration and spectral contrasts can aid L2 learners in discriminating tense-lax vowel pairs. Similarly, this study also focuses on tense-lax vowel pairs. Moreover, Yu's research also considers Bohn's hypothesis in the examination of the accuracy of participants' perceptions of vowel pairs. However, the participants are not Catalan speakers, but Mandarin L2 and Arabic L2 learners. Also, like most studies mentioned above, Yu's research does not assess training methods to the same extent as this research.

The fifth study that bears similarities with this study is Carlet and Souza's (2018) classroom-based study on the effectiveness of formal instruction of L2 vowel sounds in terms of perception and production. While they also assessed Catalan/Spanish participants and the instruction included tense-lax vowel pairs like /i/-/ɪ/, they did not consider Bohn's hypothesis in examining the performance of their participants.

1.6.4 Research questions and hypotheses

This study sets out to assess the cue weighting of duration and spectra in the perception of tense-lax vowels by naïve speakers of 13-14 years of age. To this end, there research questions have been formulated:

1. RQ1: Which acoustic cue-weighting is more significant in Catalan/Spanish vowel perception: vowel duration or spectra?

1.1 Alternative hypothesis Ha: Catalan/Spanish EFL learners will rely mainly on durational cues, instead of spectral ones, when perceiving the tense-lax vowel pair /i/-/ɪ/.

2. RQ2: To what extent do different approaches to pronunciation help the participants improve their vowel perception and production?

2.1 Alternative hypothesis Ha: Catalan-Spanish participants will improve their L2 vowel perception (i.e., they will not rely on duration to distinguish between tense and lax vowels) after training and will distinguish between tense and lax vowels in their productions.

3. RQ3: Does the quality-based training prime the participants to rely on spectral contrasts for the identification of vowels?

3.1 Alternative hypothesis Ha: subjects from the quality-based training will rely on spectral contrasts in their vowel perception after training.

Chapter 2. Methodology

The present chapter aims to provide a detailed description of the methodology adopted in both the perception and production experiments and the training instruction that the subjects received. In order to answer RQ2 and RQ3, both perception and production experiments followed a pretest-posttest design interposed by pronunciation instructions based on articulatory and acoustic descriptions of 12 English vowels, as well as contrastive analysis of L1 and L2 vowel systems.

This chapter will also describe the data collection methods, stimuli manipulations, participants, procedures, and the measurements, as well as the descriptive and statistical analysis carried out.

2.1. Experiment 1: English vowel identification task

In order to answer the research questions, a two-alternative forced-choice identification task (2AFC) was designed to elicit the cue weighting of spectra and duration in the subjects' perception of tense-lax high front vowel pairs /i/-/ɪ/. This task has been widely used by similar experiments that tested non-native discrimination of two vowel sounds close to each other in the vowel space (e.g., Bohn, 1995; Cerviño-Povedano & Mora, 2015; Escudero & Boersma, 2004; Souza et al., 2017; Wang, 2006).

2.1.1. Participants

The sample population of this study consisted of 43 bilingual Spanish-Catalan participants who studied English as a foreign language (EFL; L2 learners). The subjects were 13 and 14 years old and constituted a convenience sample, since they were students in a public school in Catalonia, where the researcher of this paper imparted classes. Importantly, the subjects were naïve speakers, as they had not received any formal instruction on English pronunciation or phonology, according to their responses to a pre-questionnaire and the lack of such instruction in the school's curriculum. Additionally, 45% of participants had never visited an English-speaking country, while 48.8% visited an English-speaking country for a period between a week and less than a month; only one participant stayed in an English-speaking country for a month.

All participants were assigned identification codes to keep track of their responses, groups, and maintain anonymity. Moreover, subjects whose L1 was English were excluded from the experiment, but were allowed to participate in the perception task even though their answers were tracked and excluded.

2.1.2. Stimuli

The forced-choice identification task consisted of a resynthesised /i/-/ɪ/ continuum that varied factorially in five equal duration and spectra steps. The resynthesised continuum yielded 25 stimuli, which were randomised and repeated in the 2AFC, resulting in a total of 50 tokens. The stimuli from the continuum were given a numerical value from one to five and coded as quality (Q) or length (L) steps (see Figure 3).

Importantly, the duration and spectra of each stimulus were manipulated in equal number of steps to compensate for the possible confounding effect of a design feature that Bohn (1995) considered could prime participants to rely on "the least varying dimension" (p. 299; section 1.3.3). Previous research has adopted this design criterion to address the claim made by Bohn (e.g., Escudero et al., 2009; Flege et al., 1997b; Yu, 2023; cf. Cebrian, 2006).

The stimuli were resynthesised from recordings of the words *beat* and *bit* by a native Canadian English speaker using a SONY PCM-M10 with a sampling rate of

44.100 Hz. With Praat software (Boersma & Weenink, 2024), the original F1, F2, and duration values of each production were calculated to carry out the relative manipulations (see Table 1). The original F1 and F2 values of /i/-/ɪ/ were calculated in Hz and established the endpoints of the spectral dimension. The difference in F1 and F2 between /i/ and /ɪ/ was calculated in Hz and divided by four steps in equal increments. Likewise, the duration of /i/ and /ɪ/ were calculated and taken as the endpoints of the duration dimension of the continuum. The difference in duration was calculated and divided by four to produce equal intermediate length values between "long" and "short" endpoints. After the vowels' acoustic dimensions were measured, the Praat Vocal Toolkit (Corretge, 2024) was used to manipulate the vowel formants and produce the intermediate stimuli. Likewise, duration was manipulated by means of duration tiers with Praat.

Figure 1 Illustration of the synthetic /i/-/ɪ/ continuum

	L1	L2	L3	L4	L5
Q1	I beat I	I Q1L2 I	I Q1L3 I	I Q1L4 I	I Q1L5 I
Q2	I Q2L1 I	I Q2L2 I	I Q2L3 I	I Q2L4 I	I Q2L5 I
Q3	I Q3L1 I	I Q3L2 I	I Q3L3 I	I Q3L4 I	I Q3L5 I
Q4	I Q4L1 I	I Q4L2 I	I Q4L3 I	I Q4L4 I	I Q4L5 I
Q5	I Q5L1 I	I Q5L2 I	I Q5L3 I	I Q5L4 I	I bit I

Note. This figure illustrates the orthogonal manipulation of spectra and duration of the stimuli. The vertical axis indicates the quality (Q), or spectrum steps and the horizontal axis indicates the length (L) or duration steps. Thus, the endpoints may be expressed as Q1L1 (*beat*) and Q5L5 (*bit*).

Table 1 Acoustic measurements of spectra and duration steps

Stimuli	Mean F1 (Hz)	Mean F2 (Hz)	Duration (ms)
Stimulus 1 *Beat*	342.83	1364.65	153.04
Stimulus 2	397.56	1496.52	137.56
Stimulus 3	452.29	1628.38	122.08
Stimulus 4	507.01	1760.25	106.60
Stimulus 5 *Bit*	561.74	1892.12	91.13
Difference	218.9	527.46	-61.91
Difference/Step	54.72	131.86	-15.47

Note. This table contains the F1, F2, and duration measurements of the original *beat-bit* tokens and those of the intermediate steps.

2.1.3. Procedure

The 50 tokens from the resynthesised continuum were presented to the subjects in a random order and aurally over earphones (DCU Tecnologic and BigBen Connected Kit Piéton). The perception task was conducted using a shared MS Forms that subjects accessed through their school's facilities. For each stimulus, participants were asked to select one of two words (*beat* and *bit*) depending on what they heard.

Whereas the same conditions applied to the post-test as the pre-test, the pre-test contained a pre-questionnaire that inquired about participants' formal pronunciation instruction and previous English linguistic experience.

2.2. Experiment 2: English tense-lax vowel production

The second experiment aimed to answer RQ3 with regards to subjects' pronunciation improvement after receiving pronunciation training. This experiment aimed to test participants' production of English tense-lax vowel pairs /i/-/ɪ/ and /u/-/ʊ/. A secondary aim of this experiment was assessing the relationship between vowel perception and production in the way participants attend to specific acoustic cues to perceive and produce vowel sounds.

2.2.1. Participants

The second experiment was conducted with a subset (N =20) of the participants from the previous experiment. The selection criterion was based on uniformity in participants' pronunciation and overall English performance as assessed by their tutor. This criterion aimed to mitigate the number of possible outliers in the test and ensure homogeneity in the sample population.

2.2.2. Procedure

Before carrying out the production experiment, 14-year-old participants were given a consent form and an information sheet (Appendix A) and so were the parents of 13-year-olds.

The production experiment was conducted with individual subjects in a quiet room in the participants' school centre. The subjects were given a sheet containing minimal pairs of tense-lax vowels /i/-/ɪ/ (/b_t/, /b_d/, /f_t/, /h_d/) and /u/-/ʊ/ (/w_d/, /p_l/, /l_k/, /s_t/) before word-final voiced and voiceless consonants, a total of 16 items. Then, the subjects were asked to pronounce the items in order and at normal speech rate. Their vowel productions were recorded using SONY PCM-M10 with a sampling rate of 44.100 Hz.

2.2.3. Acoustic measurements

The production data gathered from the subset (N =20) was coded to keep anonymity, and to track participants' responses across groups. The main acoustic dimensions (i.e., spectrum and duration) were calculated using Praat. For spectral information, the F1 and F2 midpoints of vowels were extracted (as this was considered a more stable measurement than the average value of formants) and measured in Hz. Likewise, duration was extracted and measured in milliseconds (ms). Using Praat's spectrogram and waveform representations, both formants and duration were defined manually by selecting a section of the speech, from the onset of periodicity (vocal fold vibration) to the offset of periodicity (Appendix B). This acoustic data was later extracted and compiled using a Praat script (Crosshwhite, n.d).

2.3. Pronunciation training

This study also aimed to assess the efficacy of brief, intensive pronunciation training on subjects' vowel perception and production. Thus, three pronunciation trainings were designed and administered to the subjects after the pre-test. The subjects were divided into three groups according to pronunciation training: quality-based (QT, N = 10), length-based (LT, N = 18), and implicit trainings (IT, N = 15). In total, all groups received four one-hour-long sessions of pronunciation instruction in the span of one week and a half. The production experiment also collected data from the three different groups: QT, N = 7; LT, N = 7; IT, N = 6.

All groups received the same instruction; however, vowels were taught in terms of tenseness (tense and lax vowels) to the QT group and based on length (long and short vowels) to the LT group. These methodologically different approaches aimed to assess their influence on participants' reliance on the acoustic dimensions of spectrum and duration in both perception and production.

The structure of the training sessions was partly based on the communicative framework for teaching pronunciation proposed by Farrelly (2018). The first session was dedicated to raising consciousness about the English vowel system and the lack of one-to-one equivalence between spelling and pronunciation. The second session aimed to sensitise students to the articulatory characteristics of the 12 English vowels presented and to conduct vowel discrimination activities, as well as vowel production exercises. The QT and LT groups were introduced to more specific articulatory descriptions based on tenseness and length, respectively. The third session compared L1 Catalan vowel system with L2 English vowel system by way of a contrastive analysis. It also provided further articulatory and acoustic descriptions of English sounds, as well as production and vowel perception activities. The fourth session aimed to consolidate the description, perception, and production of English monophthongs, as well as the factors of tenseness and length.

The English variety presented to the subjects in the training sessions was Southern British English (SBE), instead of American. Additionally, the phonemic transcriptions form the slides only contained length marks when shown to the TL group, whereas the QT and IT groups were not exposed to length-based classifications of vowels.

2.4. Measures and analysis

2.4.1. Perception and production data

Even though the perception task contained 50 tokens, the subjects from the IT group could not hear the stimuli Q2L2, Q3L4, Q3L3, and Q4L5 (see Figure 1) in the pre-test, due to technical issues. Therefore, statistical, and descriptive analysis excluded these stimuli in the cross-group analyses. Additionally, the analysis of the production data excluded the productions in the context /s_t/, since participants consistently mispronounced these items as diphthongs. Thus, the total number of items was reduced to 14. Moreover, outliers from the production data were excluded based on the two standard deviation criterion. That is, datapoints of F1, F2, and duration variables that were above or below the average ± 2*SD were excluded.

The subjects' responses from the 2AFC were computed by assigning numerical values from 0 to 2 to each of the stimuli per participant. The stimuli were presented twice, resulting in 50 tokens. Therefore, a maximum score of 2 indicated that a participant perceived the stimulus as *beat* in both times it was presented. Conversely, a score of 0 indicated that a participant consistently perceived the stimulus as *bit* on both occasions. And a rating of 1 indicated that a participant perceived the stimulus as *beat* and *bit* on different occasions.

Two native speakers of American English also conducted the 2AFC and their answers were compared with those of this study's sample population for valuable descriptive insights that also guide the interpretation of results.

2.4.2 Statistical analysis of perception data

To address the research questions regarding vowel perception, four statistical analyses were performed with Linear Mixed-Models (LMMs), given the complexity of the data collected.

1. *Effect of duration manipulation*: this analysis aimed to assess the effects of length manipulations on the subjects' vowel perception in the identification task. The results observed analysed in relation to the
2. *Effect of quality manipulation*: this analysis was performed to assess the degree to which subjects relied on quality in discriminating the tense-lax vowels.
3. *Estimated marginal means (EMMs) analysis*: the LMMs tested the estimated marginal means of responses given to each stimulus by each subject to compare group and time variables.
4. *Differential endpoint analysis*: this analysis afforded a more comprehensive view of the duration and quality manipulation effects. To this end, the differences between the responses between the endpoints of duration and quality were calculated and statistically tested. To calculate the effect of spectral manipulation, the response ratings (i.e., 0–2) of the last spectral stimulus endpoint (*bit*) was subtracted from the response ratings of the first spectral stimulus endpoint (*beat*) at the same length variation (axis). Thus, the only changing variable is quality, eliciting the quality effect on participants' responses. Likewise, to test the effect of duration manipulations, the response ratings of the last stimulus endpoint of the duration axis (short) were subtracted from the response ratings of the first stimulus endpoint (long). These calculations were made for each participant at pre-test and post-test and for each subset of both quality and duration axes.

2.4.3. Statistical analysis of production data

To answer RQ3, three main analyses were carried out with LMMs on the data collected from the subjects' tense-lax vowel productions.

1. *Effect of training on F1, F2, and duration*: LMMs tested significant shifts in spectra (F1 and F2) as well as duration in participants' vowel productions at the pre-test and post-test.
2. *Relative F1-F2 distance*: this analysis aimed to provide an additional view on the production performance of participants by measuring the relative distance between their F1 and F2 productions, as well as significant changes at the pre-test and post-test. The F1 of each vowel production was subtracted from their corresponding F2 value as a measure of F1-F2 distance.
3. *Relative Tense-lax distance*: a further test sought to assess the relative distance between tense and lax vowel productions at the pre-test and post-test. This was done to assess the effect of training on the discrimination of tense-lax vowels in subjects' productions. To perform this analysis, vowel productions were separated into high front and high back vowels and an average of F1, F2 and duration was extracted from the set of tokens with the same vowel quality (e.g., the average F1 of all /i/ productions in /b_t/, /b_d/, /f_t/, /h_d/ contexts per person). Within each group of vowels, tense vowel acoustic measurements were subtracted from those of lax counterparts as a measurement of tense-lax distance.

The previous tests introduced and compared native speakers' (NS) measurements of the same acoustic values measured in each test. Subjects' acoustic dimensions were statistically compared with those of NS, which effectively represented the target values. In this experiment, approximation to the NS' acoustic values (F1, F2 and duration) constituted the operational definition of vowel production improvement.

The NS' spectral and duration values selected for comparison were obtained from Hillenbrand et al. (1994). This source was considered adequate because the NS who were recorded and studied were 12 years old, approximately the same age as this study's sample population (13- 14-year-olds). Thus, the differences caused by physiological variations in vocal tract size, for example, were mitigated to a certain extent.

Chapter 3. Results

The present chapter will present the main findings from the tests described above. Incidental and noteworthy expositions of the results will be made in terms of articulatory and perceptual dimensions when reporting production and perception data.

Firstly, this chapter will expound on the observations gathered from the descriptive analysis of the perception experiment. Secondly, it will delve into the results from the statistical analyses of this data. Following this, the production experiment will be presented in terms of descriptive observations of the results before presenting the findings from the statistical analyses.

3.1. Perception experiment

In this section, the analysis and results of two-alternative forced-choice identification task are provided in detail. Briefly, this experiment set out to find out the cue weighting of vowel duration and spectra in Spanish-Catalan vowel perception (RQ1), the effect of pronunciation training on subjects' vowel perception (RQ2), and whether the quality-based training (QT) primed subjects to rely on quality contrasts (RQ3):

3.1.1. Descriptive analysis of the perception experiment

The results of the 2AFC (pre-test and post-test combined) are illustrated in Figures 3.1–3.3. The figures illustrate the function of *beat* identifications in percentages (% *beat* id.)

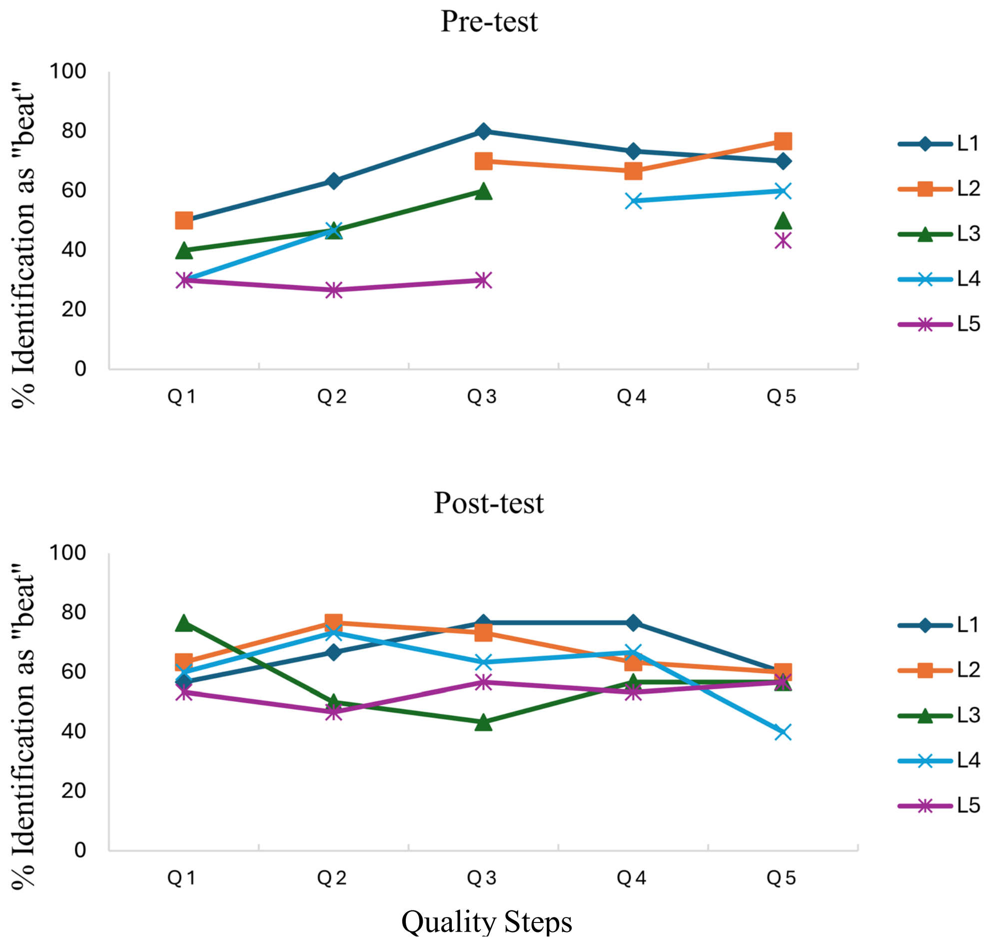
for each factorial manipulation of quality (Q) and length (L) in the continuum. The total number of responses per group and time (pre-test and post-test) have been summarised and calculated in percentages (Appendix C).

Figure 3.1 (top panel) shows the responses obtained from the IT group (N = 15) in the 2AFC in the pre-test. As can be seen in the figure, the identification functions for length variations are discrete (i.e., there are discernible separations between the *beat* responses at L1, L2, L3, L4, and L5 steps almost evenly spread along the quality axis). IT subjects consistently perceived stimuli as *beat* more often when they were long (L1), than when they were short (L5).

Likewise, Figure 3.2 (top panel) reveals a similar trend in the QT group (N = 10) as that found in Figure 3.1, albeit to a lesser extent and more inconsistently. Similarly, subjects seem to identify stimuli as *beat* more often than *bit* when long, although the difference is marginal for stimuli Q3L1 (% *beat* id. = 63.9) and Q3L5 (% *beat* id. = 55.6). Moreover, there is an increase in overlap of *beat* identifications across length manipulations.

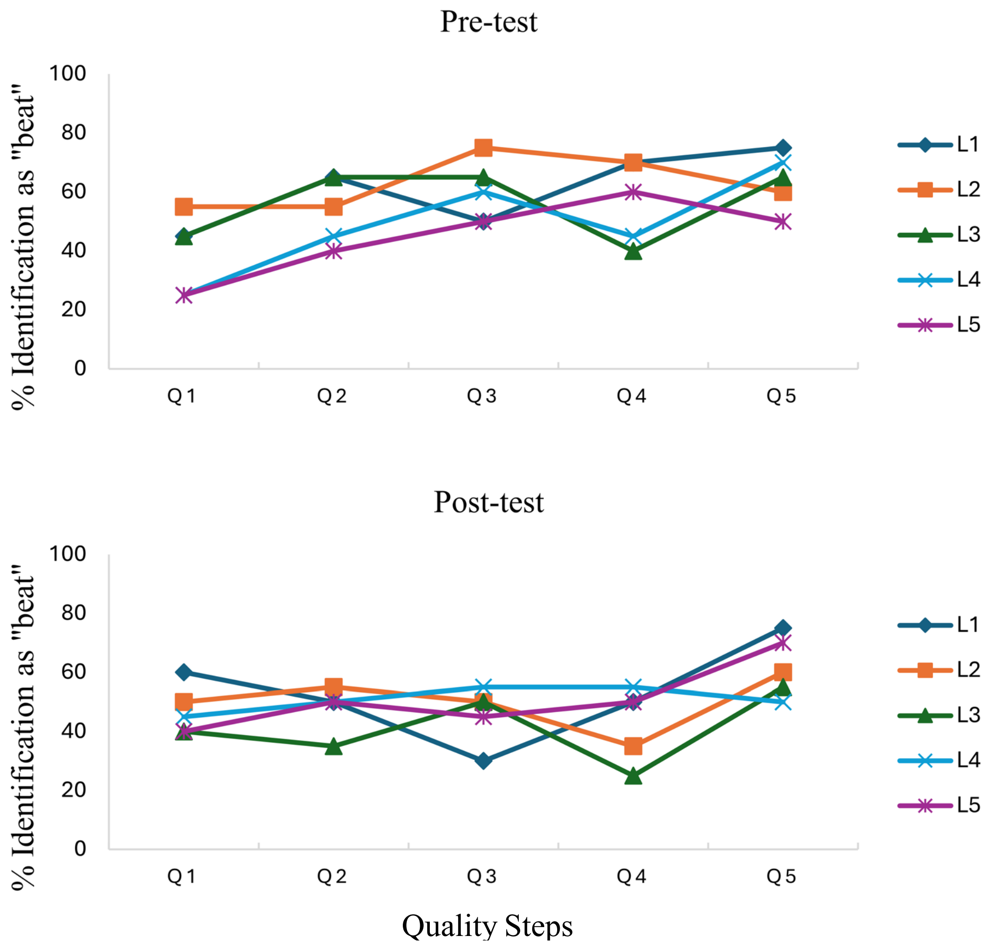
Regarding the identification scores of LT (N = 18), Table 3.3 (top panel) seems to show a higher degree of overlapping of identification functions for length variations converging in the middle of the x-axis. In other words, the concentration of *beat* identifications seems to be higher than those found in Figures 3.1 and 3.2. Moreover, the difference in percentage of *beat* identifications (% *beat* id.) between Q1L1 and Q1L5 as well as Q2L1 and Q2L5 was negative, which indicates that short stimuli for the first and second steps of the quality axis received more *beat* responses than their long counterparts. This difference was not observed in previous groups.

Figure 3. 1 IT Group's Stimuli Identification as "beat" at Pre-Test and Post-Test



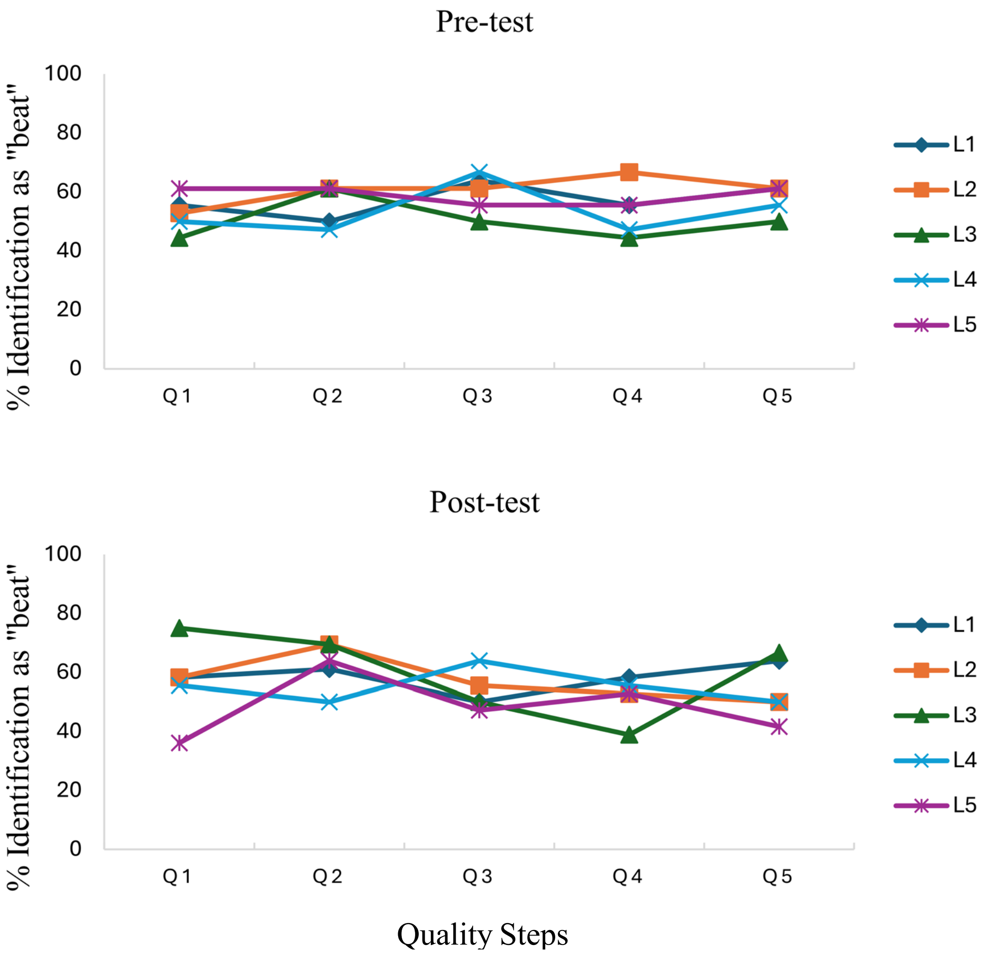
Note. This graph illustrates the identification functions for quality (Q) and length (L) steps of the continuum (see Figure 1) for the pre-test (top panel) and for the post-test (bottom panel) results.

Figure 3. 2 QT Group's Stimuli Identification as "beat" at Pre-Test and Post-Test



Note. This graph illustrates the identification functions for quality (Q) and length (L) steps of the continuum (see Figure 1) for the pre-test (top panel) and for the post-test (bottom panel) results.

Figure 3. 3 LT Group's Stimuli Identification as "beat" at Pre-Test and Post-Test



Note. This graph illustrates the identification functions for quality (Q) and length (L) steps of the continuum (see Figure 1) for the pre-test (top panel) and for the post-test (bottom panel) results.

3.1.1.1. Differences in beat identifications in pre-test and post-test

Figure 3.1 (bottom panel) demonstrates a reduction in the discrepancies between identification functions for length variations in the post-test results. Thus, the variability of participants' vowel identification across different length manipulations was less pronounced in the post-test than in the pre-test. Most noticeably, stimulus Q1 (original quality of *beat*) received more *beat* responses across all length variations than in the pre-test. Thus, more subjects correctly identified stimulus Q1 as *beat* in the post-test.

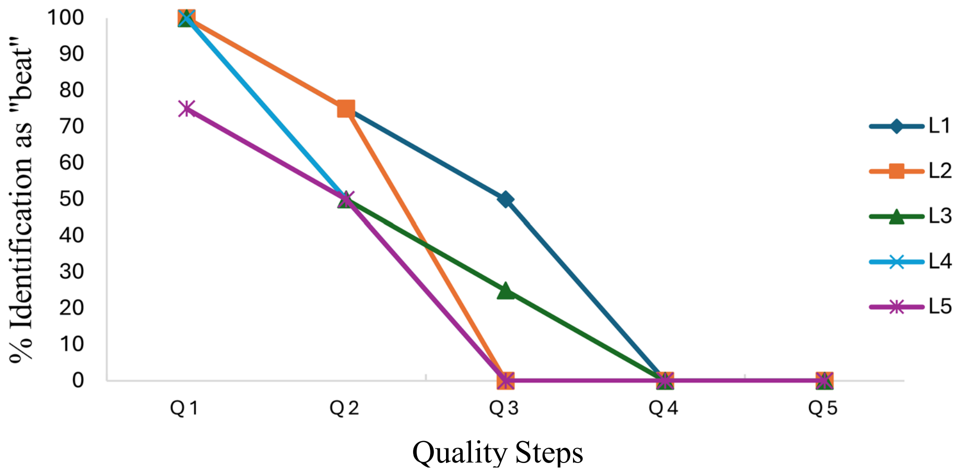
Similarly, Figure 3.2 (bottom panel) seems to show a higher degree of *beat* identification concentration across length variations in the post-test. However, differences in identification scores between different lengths can still be observed. Moreover, there is a difference in *beat* responses between long (% *beat* id. = 30) and short (% *beat* id. = 45) stimuli of Q3. Thus, the third quality step of the y-axis, a spectrally ambiguous stimulus, was more often identified as *beat* when it was short than when it was long. Noticeably, Q5 received higher *beat* identifications for both long and short (with a marginal difference of 5%) than for any other length variation.

Figure 3.3 seems to show a higher discrepancy of the percentage of *beat* identifications (% *beat* id.) across length variations in the post-test (bottom panel) than in the pre-test (top panel). However, there are fewer negative differences between long and short stimuli identifications (negative difference is understood as short stimuli having higher % *beat* id. than long ones). Only Q2 received higher number of *beat* identifications for short stimuli than for long ones. Moreover, Figure 3.3 (bottom panel) also reveals a higher positive difference in identifications between L1 and L5 at the quality endpoints Q1 and Q5. That is, long stimuli were more often identified as *beat* than short ones at both Q1 and Q5 quality endpoints.

3.1.1.2. Comparison with native speakers' vowel identification

Figure 4 shows the results from the identification task performed by two native speakers of American English. As can be seen in Figure 5, there is a steep and progressive incline in percentages of *beat* identifications (% *beat* id.) as the quality changes. Compared to Figures 3.1–3.3, Figure 4 only reveals certain disparity in vowel identifications across length variations between Q1 and Q3 stimuli. Stimuli Q4 and Q5 were only perceived as *bit* regardless of length. It is worth noting that where native speakers' responses varied, the long stimuli were more often perceived as *beat* and short stimuli as *bit*.

Figure 3. 4 Native speakers responses to identification task



Note. This graph illustrates the identification functions for quality (Q) and length (L) steps of the continuum (see Figure 3).

3.1.1.3 Averages and standard deviations across Group and Time

After describing the trends observed in the figures above, a closer analysis on the dispersion of *beat* responses and their averages was deemed adequate to assess how widespread these trends were.

Tables 2.1.1–2.3.2 contain the averages of *beat* responses ($\bar{x}$ *beat* id.) per stimulus, as well as the standard deviations (SD) of the averages. The results are colour-coded in such a way that red indicates higher average of *beat* responses (max. 2), whereas white represents higher average of *bit* responses (min. 0).

Table 2.1.1 reflects the results observed in Figure 4.1 for the IT group (N = 15) in pre-test. The gradual downward trend in average *beat* identifications ($\bar{x}$ *beat* id.) towards the short stimuli is indicated by the progressive, yet noticeable change of shade. On the other hand, quality variations seem to be consistently identified as *beat* throughout. Thus, the average of *beat* identifications seems to vary more along the length manipulations than the quality manipulations in the pre-test.

Table 2.1.1. *IT Group's Averages and Standard Deviation from Identification Task in Pre-test*

	L1		L2		L3		L4		L5	
	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Q1	1.00	0.76	1.00	0.85	0.80	0.77	0.60	0.74	0.60	0.74
Q2	1.27	0.88	–	–	0.93	0.80	0.93	0.70	0.53	0.52
Q3	1.60	0.51	1.40	0.63	1.20	0.68	–	–	0.60	0.63
Q4	1.47	0.74	1.33	0.49	–	–	1.13	0.64	–	–
Q5	1.40	0.83	1.53	0.52	1.00	0.85	1.20	0.77	0.87	0.64

Note. The table displays the synthetic continuum with the average of *beat* responses and its standard deviations per stimulus from the IT group. The results are colour-coded: red represents higher number of *beat* responses (max. 2) and white represents lower number of *beat* responses (min. 0).

Moreover, the standard deviation of the averages at the length endpoints of Q2L1 (SD = 0.88)–Q2L5 (SD = 0.52) and Q5L1 (SD = 0.83)–Q5L5 (SD = 0.64) are lower for the short variation of the stimuli (L5) than their long counterparts (L1). This indicates a lower discrepancy of identification of short stimuli than in their long counterparts for Q2 and Q5. Furthermore, it is worth noting that the most qualitatively ambiguous stimuli, those in the middle of the quality axis, seem to receive lower standard deviations than the quality endpoints (Q1 and Q5) for the same length. For example, Q3L1 (SD = 0.51) has a lower *beat* identification dispersion than Q5L1 (SD = 0.83) and Q1L1 (SD = 0.76).

Table 2.1.2, however, seems to show smaller contrasts along the duration variations. For example, participants identified L3 and L5 stimuli of Q5 as *beat* the same number of times ($\bar{x}$ *beat* id. = 0.13) despite the length variation. Moreover, it is worth noting that the absolute endpoints of the continuum (Q1L1 and Q5L5) were also identified as *beat* the same number of times (x̄ *beat* id. = 1.13), unlike the results of the pre-test (Table 2.1.1, Q1L1, x̄ *beat* id. = 1; Q5L5, x̄ *beat* id. = 0.87).

Table 2.1.2. IT Average and Standard Deviation from Identification Task in Post-test

	L1		L2		L3		L4		L5	
	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Q1	1.13	0.83	1.27	0.80	1.53	0.74	1.20	0.86	1.07	0.96
Q2	1.33	0.72	1.53	0.52	1.00	0.76	1.47	0.74	0.93	1.03
Q3	1.53	0.64	1.47	0.64	0.87	0.92	1.27	0.80	1.13	0.92
Q4	1.53	0.74	1.27	0.70	1.13	0.83	1.33	0.62	1.07	0.80
Q5	1.20	0.94	1.20	0.94	1.13	0.92	0.80	0.68	1.13	0.92

Note. The table displays the synthetic continuum with the average of *beat* responses and standard deviations per stimulus from the IT group. The results are colour-coded: red represents higher number of *beat* responses (max. 2) and white represents lower number of *beat* responses (min. 0).

Furthermore, Table 2.1.2 (post-test) shows how the short stimuli (L5) of all quality variations consistently exhibited higher standard deviations in *beat* responses than in Figure 2.1.1 (pre-test). This indicates a higher dispersion of *beat* responses for short stimuli in the post-test. Similarly, the average of *beat* identifications of short stimuli also increased in the post-test.

Table 2.2.1 seems to reveal a similar trend for the QT group (N = 10) as that found in Tables 2.1.1 and 2.1.2 because shorter stimuli appear to receive lower *beat* responses than longer ones, albeit to a lesser degree. For example, subjects identified long stimuli Q5L1 as *beat* more often (x̄ *beat* id. = 1.50) than their short counterparts (x̄ *beat* id. = 1). Even though the gradated incline of *beat* responses across length variations is less pronounced in the table below (cf. Table 2.1.1), the stark contrasts between the long and short endpoints still exist. All short stimuli are consistently perceived more often as *bit* than their long counterparts, except Q3L1–Q3L5: both had an average of *beat* identification of 1 (i.e., half the participants identified them as *beat*).

Nevertheless, the standard deviation of the duration endpoints seems to be more homogeneous than that found in Table 2.1.1. The table below shows less dispersion of *beat* responses at some, but not all the short stimuli. For example, Q1L5 (SD = 0.53) has less disparity of responses than Q1L1 (SD = 0.88) and the endpoints Q5L1 (SD = 0.71) and Q5L5 (SD = 0.67) mirror the same relation of dispersion at the endpoints with answers for the shorter stimuli being more concentrated. On the other hand, ambiguous values in the middle of the continuum have higher standard deviations than their long counterparts except for Q3L1 (SD = 0.67) – Q3L5 (SD = 0.67).

However, in the post-test (Table 2.2.2) QT subjects' perception of the stimuli seem to have evened out along the continuum and settled around the midpoint of average *beat* identifications (i.e., x̄ *beat* id. = 1). This means that *beat* identifications per stimulus were lower in the post-test on average throughout the continuum. Moreover, the maximum standard deviation observed in the pre-test increased (SD = 0.94) compared to the post-test (SD = 88), as can be seen in stimuli Q2L1, Q4L1, Q1L2, Q5L4, and Q2L5, indicating higher variation in perceptions of these stimuli.

Table 2.2.1 QT Group's Averages and Standard Deviation from Identification Task in Pre-test

	L1		L2		L3		L4		L5	
	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Q1	0.90	0.88	1.10	0.57	0.90	0.74	0.50	0.71	0.50	0.53
Q2	1.30	0.67	1.10	0.88	1.30	0.67	0.90	0.88	0.80	0.79
Q3	1.00	0.67	1.50	0.71	1.30	0.82	1.20	0.63	1.00	0.67
Q4	1.40	0.52	1.40	0.52	0.80	0.79	0.90	0.74	1.20	0.79
Q5	1.50	0.71	1.20	0.63	1.30	0.82	1.40	0.70	1.00	0.67

Note. The table displays the synthetic continuum with the average of *beat* responses and standard deviations per stimulus from the QT group. The results are colour-coded: red represents higher number of *beat* responses (max. 2) and white represents lower number of *beat* responses (min. 0).

Table 2.2.2 **QT Group's Averages and Standard Deviation from Identification Task in Post-test**

	L1		L2		L3		L4		L5	
	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Q1	1.20	0.79	1.00	0.94	0.80	0.79	0.90	0.57	0.80	0.79
Q2	1.00	0.94	1.10	0.74	0.70	0.82	1.00	0.82	1.00	0.94
Q3	0.60	0.84	1.00	0.67	1.00	0.67	1.10	0.88	0.90	0.74
Q4	1.00	0.94	0.70	0.67	0.50	0.53	1.10	0.74	1.00	0.82
Q5	1.50	0.71	1.20	0.79	1.10	0.74	1.00	0.94	1.40	0.84

Note. The table displays the synthetic continuum with the average of *beat* responses and standard deviations per stimulus from the QT group. The results are colour-coded: red represents higher number of *beat* responses (max. 2) and white represents lower number of *beat* responses (min. 0).

Table 2.3.1 represents the results of the identification task from the LT group (N = 18). While the previous results (Tables 2.1.1–2.2.2) illustrated a progressive incline in *beat* perceptions towards the endpoint of the duration axis, Table 2.3.1 illustrates a concave pattern in perceptions. That is, the endpoints of the length axis seems to concentrate higher values of *beat* identifications, while the middle section of the continuum (L3 and L4) gathered fewer *beat* identifications. The centre-most stimuli are the most ambiguous stimuli since they are halfway between each endpoint and subjects seem to perceive them as *bit* (min. $\bar{x}$ *beat* id. = 0.89) more often than *beat*.

Regarding the standard deviations, high variability in responses to stimuli seem to appear in duration axes L3 and L4, the mid-section of the continuum, especially for stimuli Q1L3 (SD = 0.90), Q1L4 (SD = 0.91) and Q5L4 (SD = 0.90).

Table 2.3.1 LT Group's Averages and Standard Deviation from Identification Task in Pre-test

	L1		L2		L3		L4		L5	
	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Q1	1.11	0.68	1.06	0.73	0.89	0.90	1.00	0.91	1.22	0.65
Q2	1.00	0.77	1.22	0.65	1.22	0.73	0.94	0.80	1.22	0.81
Q3	1.28	0.46	1.22	0.81	1.00	0.77	1.33	0.69	1.11	0.68
Q4	1.11	0.76	1.33	0.77	0.89	0.83	0.94	0.87	1.11	0.76
Q5	1.22	0.81	1.22	0.73	1.00	0.77	1.11	0.90	1.22	0.73

Note. The table displays the synthetic continuum with the average of *beat* responses and standard deviations per stimulus from the LT group. The results are colour-coded: red represents higher number of *beat* responses (max. 2) and white represents lower number of *beat* responses (min. 0).

The post-test (Table 2.3.2) of LT, however, seems to mirror the trend observed in post-tests of other groups, where *beat* perceptions seem to be more evenly spread across the continuum. Moreover, a trend that was not observed in the pre-test of LT (Table 2.3.1) is that short stimuli had fewer *beat* identifications, aligning with the previous patterns observed in Tables 2.1.1–2.2.2. For example, the endpoint Q1L1 (Table 2.3.2, $\bar{x}$ *beat* id. = 1.17) was perceived as *beat* more often than its short counterpart Q1L5 ($\bar{x}$ *beat* id. = 0.72). Likewise, the stimulus Q5L1 ($\bar{x}$ *beat* id. = 1.28) was more often perceived as *beat* than Q5L5 ($\bar{x}$ *beat* id. = 0.83). An exception to this pattern is the stimuli Q2L1 ($\bar{x}$ *beat* id. = 1.22) and Q2L5 ($\bar{x}$ *beat* id. = 1.28) with a marginal difference of 0.05 in their averages of *beat* identifications.

Regarding the high standard deviations observed in the stimuli Q1L3 (SD = 0.90), Q1L4 (SD = 0.91) and Q5L4 (SD = 0.90) in the pre-test (Table 2.3.1) decreased in the post-test to an extent (Table 2.3.2): Q1L3 (SD = 0.62), Q1L4 (SD = 0.76), and Q5L4 (SD = 0.84). The higher degree of variability of answers seems to have shifted to the short endpoint of the duration continuum.

Table 2.3.2 LT Group's Averages and Standard Deviation from Identification Task in Post-test

	L1		L2		L3		L4		L5	
	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Q1	1.17	0.79	1.17	0.79	1.50	0.62	1.11	0.76	0.72	0.89
Q2	1.22	0.81	1.39	0.70	1.39	0.70	1.00	0.59	1.28	0.83
Q3	1.00	0.84	1.11	0.68	1.00	0.84	1.28	0.75	0.94	0.80
Q4	1.17	0.71	1.06	0.80	0.78	0.73	1.11	0.83	1.06	0.80
Q5	1.28	0.75	1.00	0.91	1.33	0.84	1.00	0.84	0.83	0.79

Note. The table displays the synthetic continuum with the average of *beat* responses and standard deviations per stimulus from the LT group. The results are colour-coded: red represents higher number of *beat* responses (max. 2) and white represents lower number of *beat* responses (min. 0).

Overall, the results as summarised in Tables 2.1.1–2.3.2 seem to reveal a pattern to different extents across the groups, but perhaps best reflected in the IT group (Tables 2.1.1 and 2.1.2). This pattern seems to suggest that shorter stimuli received fewer *beat* responses than longer ones. Moreover, relatively high standard deviations have been observed throughout the subjects' average *beat* identifications across groups.

It is worth noting that an additional pattern has been observed throughout all the groups' results with regard the identification of endpoint stimuli. All the groups, in pre-test and post-test, have consistently perceived the quality endpoint Q5, at both L1 and L5, as *beat* more often than quality endpoint Q1 at L1 and L5 (except for the LT's perception of Q5L1–L5 in the pre-test, which obtained the same average of *beat* identifications). This finding is relevant because the spectral make-up of Q1 is *beat* (/i/), while that of Q5 is *bit* (/ɪ/). Therefore, subjects tended to identify the opposite vowel quality.

3.1.1.4 Differential analysis of vowel identification

The previous section reported some patterns in vowel perceptions between pre-test and post-test. Table 3 shows some of these patterns in differential scores between pre-test and post-test. The difference was obtained by subtracting the average *beat* identification scores in the pre-test from that of the post-test. Red and white values indicate higher and lower *beat* identifications, respectively. Moreover, negative averages have been highlighted in blue.

Table 3 shows an increase of *beat* identifications towards the end of the length continuum of Q1 (from L3 to L5) as indicated by the darker shades of red. This suggests that subjects identified the shorter Q1 stimuli as *beat* more often in the post-test. However, the opposite is true of the LT group for stimulus Q1L5, which they perceived as *beat* less often in the post-test. Likewise, the QT also dropped their average *beat* identification scores for stimuli Q1L2 and Q1L3.

Considering that the stimuli in quality axis Q1 maintained the original spectral characteristics of the tense high front vowel, higher number of *beat* identifications indicated more correct answers (see Figure 5). Conversely, stimuli in quality axis Q5 had the spectral make-up of the high front lax vowel; thus, negative values indicated more correct answers, as stimuli were correctly identified as *bit* more often. This was the case for most length variations of Q5 stimuli, where most stimuli across groups were perceived as *bit* more frequently in the post-test than in the pre-test (Table 3).

Moreover, it is worth noting that the QT group's *beat* identifications for all length variations of Q3 decreased, as indicated by the negative values (Table 3). This indicates that the most qualitatively ambiguous stimuli were perceived more often as *bit* in the post-test. This change, however, seems to be more prominent in the longer stimuli of the Q3 axis with a difference of -0.40 average *beat* identifications for L1 and of -0.10 for L5.

Table 3 Difference in vowel identifications between pre-test and post-test across groups

Q1	Groups	L1	L2	L3	L4	L5
	IT	0.13	0.27	0.73	0.60	0.47
	QT	0.30	-0.10	-0.10	0.40	0.30
Q2	LT	0.07	0.11	0.61	0.11	-0.50
	IT	0.07		0.07	0.53	0.40
	QT	-0.30	0.00	-0.60	0.10	0.20
Q3	LT	0.22	0.17	0.17	0.06	0.06
	IT	-0.07	0.07	-0.33		0.53
	QT	-0.40	-0.50	-0.30	-0.10	-0.10
Q4	LT	-0.28	-0.11	0.00	-0.06	-0.17
	IT	0.07	-0.07		0.20	
	QT	-0.40	-0.70	-0.30	0.20	-0.20
Q5	LT	0.06	-0.28	-0.11	0.17	-0.06
	IT	-0.20	-0.33	0.13	-0.40	0.27
	QT	0.00	0.00	-0.20	-0.40	0.40
	LT	0.06	-0.22	0.33	-0.11	-0.39

Note. This table displays the difference in the average of *beat* identifications per stimulus between pre-test and post-test. The average of *beat* identifications for pre-test was subtracted from that of the post-test. The results are colour-coded: red represents higher number of *beat* responses and white represents lower number of *beat* responses. Negative values are highlighted in blue.

3.1.2 Statistical Analyses

As indicated in section 2.4.2, four statistical analyses were conducted on the results of the identification task to find out 1) the effect of duration manipulation on subjects' vowel perception, 2) the effect of quality manipulation on subject's vowel perception, 3) estimated marginal means across *groups*, and *time* (pre-test and post-test), 4) differential analysis of perception scores across endpoints. The results gathered from the analyses will be reported in the same order in which they were introduced.

3.1.2.1 Effect of duration manipulation on vowel quality perception

3.1.2.1.1 Effects on Q1

A Linear Mixed Models (LMMs) was conducted on the identification scores of subjects across independent variables of *group* (IT, QT, and LT), *time* (pre-test and post-test), and *length* (L1, L2, L3, L4, and L5). Table 4.1 contains the results of the LMMs on the effects of *length, group,* and *time* on the vowel identification of stimuli Q1. The analysis found *time* ($F(1, 40.05)= 4.891$, $p= .033$) and *length* ($F(4, 67.46)= 2.892$, $p= .029$) to be statistically significant on the perception of vowel qualities; however, there were not significant differences in vowel identifications across *group*. Thus, the main effects of *time* and *length* were found to be influencing the three groups equally.

Thus, vowel length proved to be a significant factor in the way subjects identified the quality of vowels in the identification task. Likewise, the pronunciation training also influenced student's vowel identification; however, the nature and orientation of these changes can be further clarified by examining the estimated marginal means (Table 4.1.2).

Table 4.1.1 Results of Linear Mixed Models on length, group, and time effects on Q1

Effect	**df**	**F**	**p**
Group	2, 40.06	1.310	0.281
Time	1, 40.05	4.891	0.033
Length	4, 67.46	2.892	0.029
Group ✻ Time	2, 40.05	1.332	0.275
Group ✻ Length	8, 67.46	0.347	0.944
Time ✻ Length	4, 240.03	1.218	0.304
Group ✻ Time ✻ Length	8, 240.03	1.677	0.105

Table 4.1.2 *Estimated Marginal Means for Q1 identification across time and group*

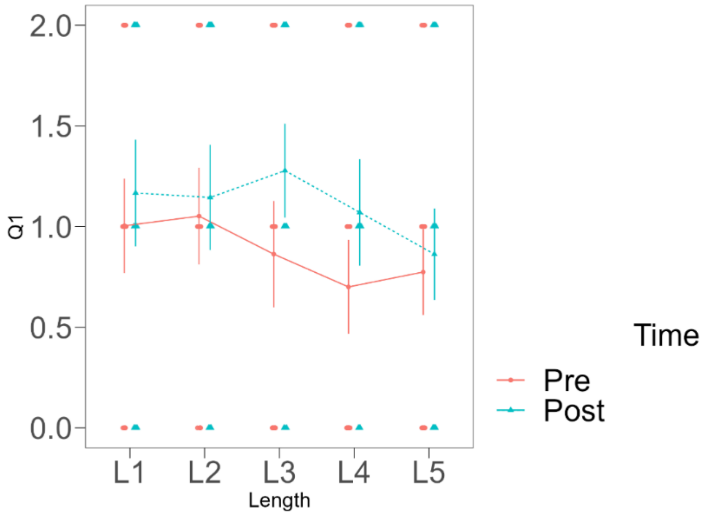
				95% CI	
Time	**Group**	**Estimate**	**SE**	**Lower**	**Upper**
Pre	Implicit	0.800	0.118	0.569	1.031
Post	Implicit	1.240	0.135	0.976	1.504
Pre	Length	1.056	0.108	0.844	1.267
Post	Length	1.133	0.123	0.892	1.375
Pre	Quality	0.780	0.145	0.497	1.063
Post	Quality	0.940	0.165	0.616	1.264

Note. This table includes the results averaged over the levels of *Length.*

As can be seen in the Table 4.1.2, the estimated marginal means (EMMs) for post-test are consistently higher than those of the pre-test for all groups. Therefore, the significant effect of *time* affected subjects in such a way that they correctly perceived the tokens (Q1) as /i/ more often in the post-test than in the pre-test.

Additionally, Figure 6.1.1 demonstrates how *length* and *time* influenced subjects' vowel identifications in the pre-test and post-test. In the figure, a separation can be observed between the identification scores of pre-test and post-test. However, consistent slopes can also be seen, indicating lower *beat* identifications as the vowel length shortens. Table 4.1.3 further illustrates this trend in EMMs across *time* and *length,* where all the post-test Q1 stimuli have higher EMMs across the length manipulations than in the pre-test (e.g., pre-test Q1L1 EMM = 1.004 vs. post-test Q1L1 EMM = 1.167). Nevertheless, a consistent decrease in vowel identifications can be observed in the pre-test and post-test (e.g., pre-test Q1L1, EMM = 1.004; Q1L5, EMM = 0.774 vs. post-test Q1L1, EMM = 1.167; Q1L5, EMM = 0.863).

Figure 6.1.1 Plot of vowel identification of Q1 as a function of length and time



Note. The figure depicts the vowel identifications in EMMs for Q1 across all length manipulations (L) in the pre-test and post-test. Results averaged over levels of *group*.

Table 4.1.3 Estimated Marginal Means for Q1 identification across length and time

				95% CI	
Length	**Time**	**Estimate**	**SE**	**Lower**	**Upper**
L1	Pre	1.004	0.118	0.773	1.234
L2	Pre	1.052	0.120	0.816	1.288
L3	Pre	0.863	0.132	0.604	1.122
L4	Pre	0.700	0.117	0.470	0.930
L5	Pre	0.774	0.107	0.563	0.985
L1	Post	1.167	0.133	0.907	1.427
L2	Post	1.144	0.131	0.888	1.401
L3	Post	1.278	0.117	1.048	1.507
L4	Post	1.070	0.132	0.811	1.330

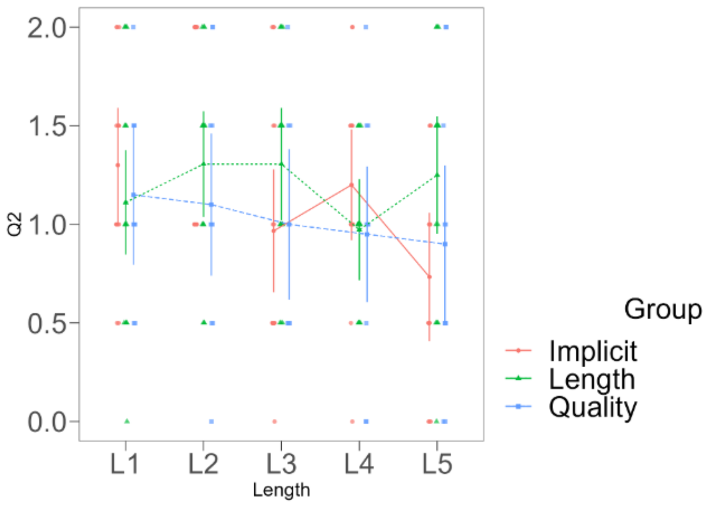
Length	Time	Estimate	SE	95% CI	
				Lower	Upper
L5	Post	0.863	0.114	0.639	1.087

Note. The results are averaged over the levels of *Group*.

3.1.2.1.2 Effects on Q2

There are no significant effect on vowel perception for Q2 from *group, time,* or *length,* as no variable obtained a p-value < .05. Thus, vowel identifications of Q2 stimuli were not affected by length manipulations or the pronunciation training. Moreover, no significant difference among groups was found to be significant. Figure 6.2 illustrates the vowel identification of the three groups for Q2 across all vowel lengths. While the IT and LT groups' identifications seem to vary greatly in different length variations, the QT's identifications seem to be more stable. Nevertheless, no significance was found for the effect of *length* (p = 0.226) or *group* (p = 0.450).

Figure 6.2 Plot of vowel identification as a function of group and length



3.1.2.1.3 Effects on Q3

Regarding Q3 perception, the LMMs returned significant effects for *length* (F(4, 60.16)= 3.321, p= 0.16); however, neither *group* (p = .406) nor *time* (p = .153) were found to be statistically significant. Thus, in a similar way to Q1 (Table 4.1.1), length manipulations influenced participants' vowel identifications equally across the three groups and both pre-test and post-test.

Table 4.3.1 Results of LMMs on length, group, and time effects on Q3

Effect	df	F	p
Group	2, 40.94	0.923	0.406
Time	1, 44.77	2.118	0.153
Length	4, 60.16	3.321	0.016
Group ＊ Time	2, 46.23	1.541	0.225
Group ＊ Length	8, 62.73	1.658	0.127
Time ＊ Length	4, 227.69	1.013	0.401
Group ＊ Time ＊ Length	7, 227.69	0.848	0.549

While *length* was significant, its precise effect on vowel identification is not as clear as that of Q1 (Table 4.1.2). Table 4.3.2 shows the EMMs of vowel identifications across pre-test and post-test as well as length variations. While the pre-test's EMMs seem to point to a gradual decline in *beat* vowel identifications as the length decreases, this trend is not as clear for the post-test, as some shorter stimuli received more *beat* identifications than longer ones (e.g., Q3L4, EMM = 1.215 vs. Q3L1, EMM = 1.044).

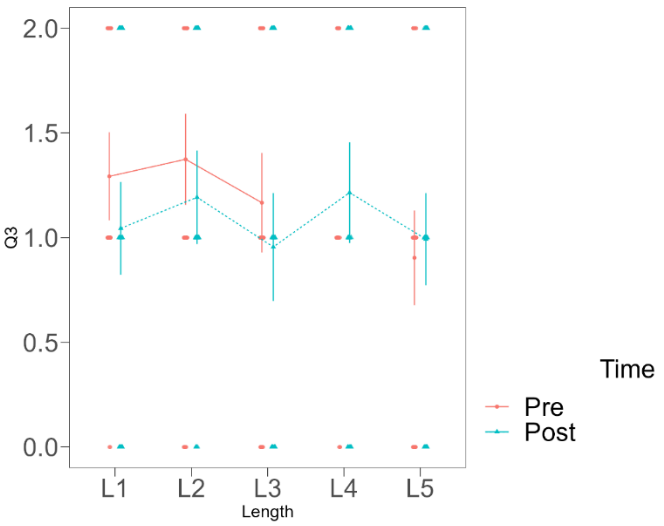
Table 4.3.2 Estimated Marginal Means for Q3 identification across length and time

Length	Time	Estimate	SE	95% CI	
				Lower	Upper
L1	Pre	1.293	0.106	1.085	1.500
L2	Pre	1.374	0.109	1.160	1.588
L3	Pre	1.167	0.119	0.933	1.400
L4	Pre				
L5	Pre	0.904	0.114	0.681	1.126
L1	Post	1.044	0.111	0.826	1.262
L2	Post	1.193	0.112	0.973	1.412
L3	Post	0.956	0.129	0.703	1.209
L4	Post	1.215	0.120	0.979	1.451
L5	Post	0.993	0.111	0.776	1.209

Note. This table includes the results averaged over the levels of *Group*.

Figure 6.3 illustrates the EMMs of *beat* vowel identifications across *length* and *time*. This figure reveals an irregular, jagged pattern of vowel identifications for post-test, where vowels were more often identified as *beat* at intermediate duration steps (L2, L4). However, the vowel identifications from pre-test can be seen to progressively decline as length variations are shorter.

Figure 6.3 Plot of vowel identification as a function of time and length



3.1.2.1.4 Effects on Q4

According to Table 4.4.1, the perception of stimuli Q4 was found to be statistically affected by *length* (F(4, 56.52)= 3.454, p= 0.14), while *time* (p = .268) and *group* (p = .204) were not statistically significant. Table 4.4.2 further shows the relationship between length variation and vowel identification. However, the influence of *length* on vowel identifications is not as regular and gradual as that on Q1 stimuli. The EMMs across the length variations in the post-test seem to vary in an irregular pattern where the short endpoint has higher *beat* identifications (L5, EMM = 1.041) than the second length variation (L2, EMM = 1.007). Nevertheless, long endpoints still received higher *beat* identifications than short endpoints both in pre-test and post-test.

Table 4.4.1 Results of LMMs on length, group, and time effects on Q4

Effect	df	F	p
Group	2, 40.53	1.655	0.204
Time	1, 46.12	1.258	0.268
Length	4, 56.52	3.454	0.014
Group ✻ Time	2, 48.28	1.240	0.298
Group ✻ Length	8, 61.10	0.468	0.874
Time ✻ Length	4, 218.95	1.917	0.109
Group ✻ Time ✻ Length	6, 218.95	0.353	0.908

Table 4.4.2 Estimated Marginal Means for Q4 identification across length and time

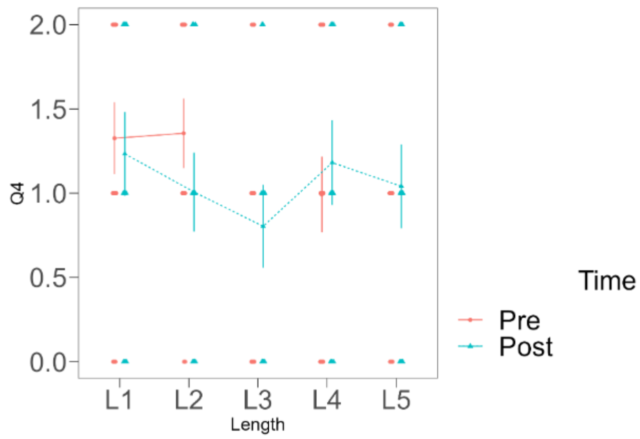
				95% CI	
Length	**Time**	**Estimate**	**SE**	**Lower**	**Upper**
L1	Pre	1.326	0.108	1.115	1.537
L2	Pre	1.356	0.104	1.153	1.559
L3	Pre				
L4	Pre	0.993	0.113	0.771	1.214
L5	Pre				
L1	Post	1.233	0.124	0.990	1.476
L2	Post	1.007	0.118	0.777	1.238

				95% CI	
Length	**Time**	**Estimate**	**SE**	**Lower**	**Upper**
L3	Post	0.804	0.123	0.562	1.045
L4	Post	1.181	0.126	0.935	1.428
L5	Post	1.041	0.124	0.797	1.285

Note. This table includes the results averaged over the levels of *Group.*

Furthermore, Figure 6.4, when compared to Figure 6.3 seems to mirror the jagged shape of the vowel identification EMMs across duration stimuli, albeit to a lesser degree. It is worth noting that the mid-point of the length continuum received the least *beat* identifications (post-test, Q4L3, EMM = 0.804).

Figure 6.4 Plot of vowel identification of Q4 as a function of time and length



3.1.2.1.5 Effects for Q5

There are no variable was statistically significant in the subjects' vowel identification of Q5 stimuli. Subjects were not influenced by *length* (p = .154), *time* (p = .534) or *group* (p = .641) when identifying Q5 in different duration steps. As can be seen in Figure 6.5, a similar jagged-shaped seems to emerge from the EMMs of vowel identifications. A similar shape was observed in Figures 6.3 and 6.4, which demonstrated irregular patterns of vowel identifications across vowel lengths of different qualities.

Table 4.5.2 shows the EMMs of vowel identifications of stimuli Q5. Even though there is no statistically significant effect of *length* on vowel perception, long stimuli are invariably perceived as *beat* more often than short stimuli (pre-test, Q5L1, EMM = 1.374; Q5L5, EMM = 1.030; post-test, Q5L1, EMM = 1.326, Q5L5, EMM = 1.122).

Figure 6.5 Plot of vowel identification of Q5 as a function of time and length

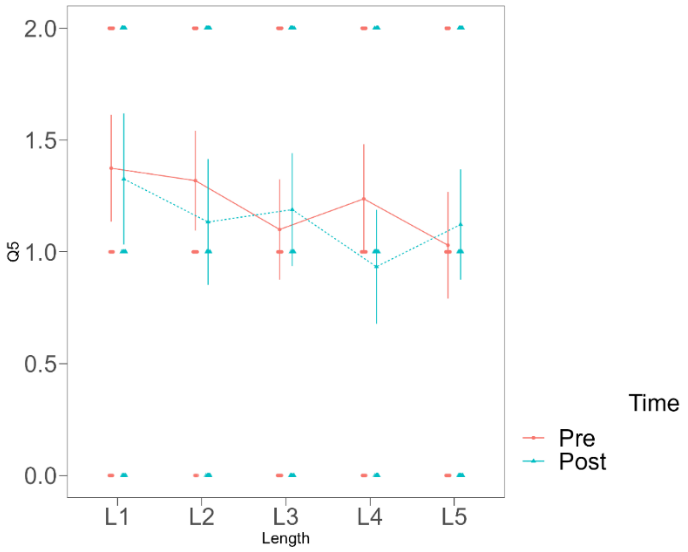


Table 4.5.2. Estimated Marginal Means for vowel identification across length and time

				95% CI	
Length	Time	Estimate	SE	Lower	Upper
L1	Pre	1.374	0.120	1.139	1.609
L2	Pre	1.319	0.112	1.099	1.538
L3	Pre	1.100	0.112	0.880	1.320
L4	Pre	1.237	0.122	0.997	1.477
L5	Pre	1.030	0.120	0.795	1.264
L1	Post	1.326	0.146	1.039	1.613
L2	Post	1.133	0.141	0.858	1.409
L3	Post	1.189	0.127	0.941	1.437
L4	Post	0.933	0.128	0.683	1.184

				95% CI	
Length	**Time**	**Estimate**	**SE**	**Lower**	**Upper**
L5	Post	1.122	0.123	0.880	1.364

Note. The results are averaged over the levels of *group*.

3.1.2.2 Effect of quality manipulation on vowel perception

Another Linear Mixed Models (LMMs) was run on the vowel identifications across length variations with independent variables of *group* (IT, QT, and LT), *time* (pre-test and post-test), and *quality* (Q1, Q2, Q3, Q4, and Q5).

3.1.2.2.1 Effects on L1

The effects of *group* (p = .093), *time* (p = .598), and *quality* (p = .456) were found not to be statistically significant in the vowel identification of long stimuli. Therefore, subjects across the three groups and pre-test and post-test where not significantly affected by the different qualities of the long stimuli when identifying them (Q1L1–Q5L1). However, Table 4.6.2 shows a somewhat gradual increase in *beat* vowel identifications as the stimuli approach the Q5 endpoint at an average increase interval of 0.066. Noticeably, the Q1L1 was identified as *bit* more often (EMM = 1.085) than Q5L1 (EMM = 1.350).

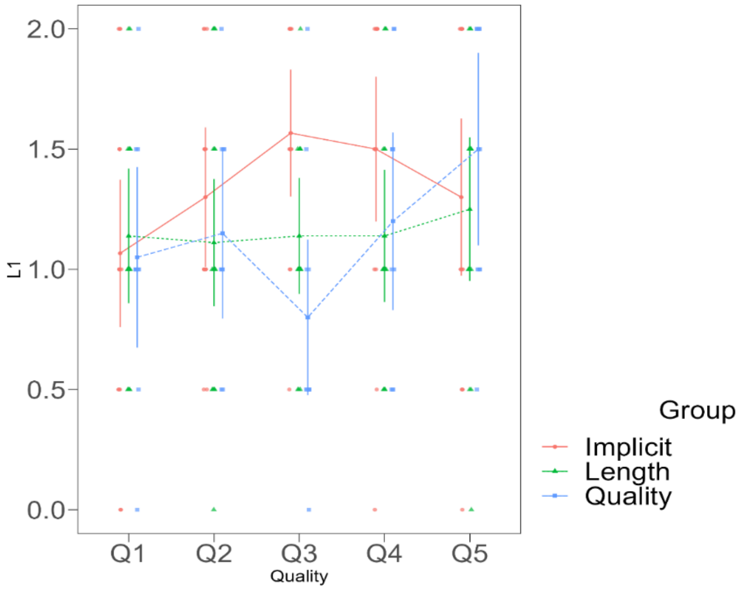
Table 4.6.2 Estimated Marginal Means for vowel identification across qualities

			95% CI	
Quality	**Estimate**	**SE**	**Lower**	**Upper**
Q1	1.085	0.092	0.904	1.266
Q2	1.187	0.087	1.016	1.358
Q3	1.169	0.080	1.013	1.324
Q4	1.280	0.091	1.102	1.458
Q5	1.350	0.098	1.157	1.543

Note. Results are averaged over the levels of *group* and *time*.

Moreover, Figure 6.6 also reveals a disparity between QT and IT groups as regards vowel identifications of Q3L1, even though *group* was not found to be statistically significant. On the other hand, the IT group (EMM = 1.067), LT (EMM = 1.139) and QT (EMM = 1.050) seem to converge in the *beat* identification of Q1L1.

Figure 6.6. Plot of vowel identification as a function of group and quality



3.1.2.2.2 Effects for L2

The LMMs did not find *quality* (p = .449), *group* (p = .174), or *time* (p = .193) statistically significant in subject's vowel perceptions of stimuli with L2 duration variation. Thus, subjects' vowel identifications did not vary significantly along the quality continuum. It is worth noting, however, that Q1 received on average lower *beat* identifications (EMM = 1.098) than its counterpart Q5 (EMM = 1.226), as indicated in Table 4.7.2. A similar difference was observed in the perception of L1 stimuli.

Table 4.7.2. Estimated Marginal Means for vowel identification across qualities

			95% CI	
Quality	**Estimate**	**SE**	**Lower**	**Upper**
Q1	1.098	0.096	0.910	1.286
Q2				
Q3	1.283	0.079	1.128	1.439
Q4	1.181	0.082	1.022	1.341
Q5	1.226	0.089	1.052	1.400

Note. Results are averaged over the levels of *group* and *time*.

Figure 6.7. Plot of vowel identification as a function of time and quality

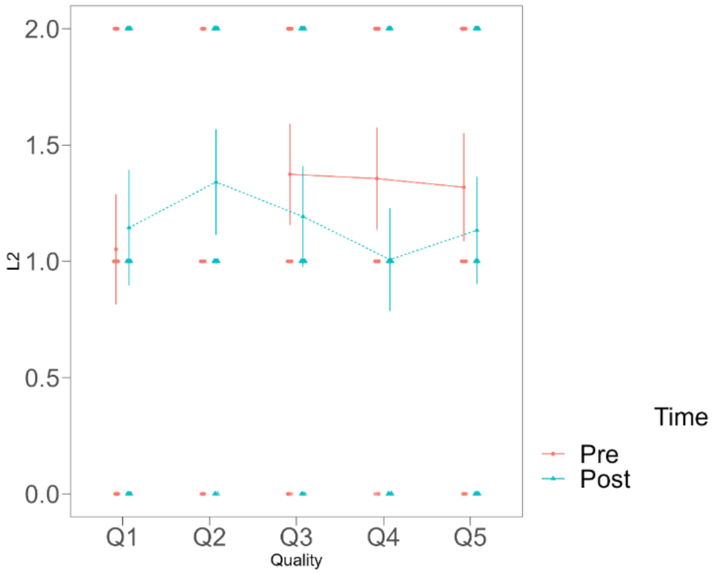


Figure 6.7 illustrates the EMMs of identifications of different vowel qualities of equal length. The figure reveals that the lower quality steps (Q3, Q4, and Q5) received fewer *beat* identifications in the post-test than in the pre-test. However, the quality endpoint Q5 was more often identified as *beat* than Q4.

3.1.2.2.3 Effects on L3

While Table 4.8.1 shows the LMMs did not find significant main effects for *group, time*, or *quality* on L3 stimuli, it encountered a significant interaction between *time* and *quality* (F(4,226.59)= 2.794, p= 0.027). This interaction is further clarified in Figure 6.8 which illustrates the change in vowel identification across *time*. While in the pre-test there is an initial increase in *beat* vowel identification from Q1 to Q3, the post-test's vowel identifications display an incline that descends from Q1 to Q4. However, stimulus Q5 received more *beat* identifications (EMM = 1.189) than other qualitatively different stimuli other than Q1 (EMM = 1.278).

Table 4.8.1. Results of LMMs on quality, group, and time effects on L3

Effect	df	F	p
Group	2, 40.93	0.540	0.587
Time	1, 44.01	0.004	0.947
Quality	4, 63.86	1.759	0.148
Group ✻ Time	2, 45.35	2.742	0.075
Group ✻ Quality	8, 65.23	1.120	0.362
Time ✻ Quality	4, 226.59	2.794	0.027
Group ✻ Time ✻ Quality	7, 226.59	0.559	0.788

Figure 6.8. Plot of vowel identification as a function of time and quality

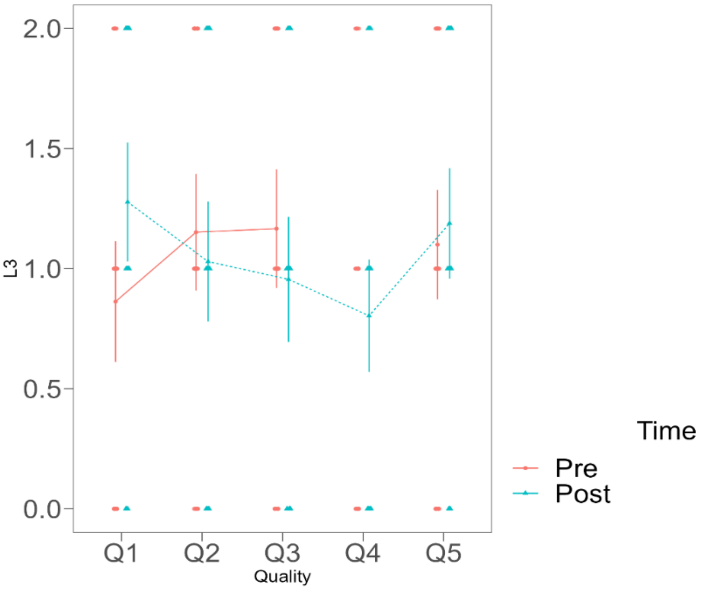


Table 4.8.2 further shows the difference in vowel identification between pre-test and post-test. The stimuli Q1–Q3 present an upward trend in identifications (EMM = 0.863–1.167), while those of the post-test exhibit a downward trend for stimuli Q1–Q4 (EMM = 1.100–0.804). Moreover, in the post-test Q1 was correctly identified as *beat* more often (EMM = 1.278) than in the pre-test (EMM = 0.863).

Table 4.8.2. Estimated Marginal Means for vowel identification across qualities

				95% CI	
Quality	**Time**	**Estimate**	**SE**	**Lower**	**Upper**
Q1	Pre	0.863	0.126	0.616	1.110
Q2	Pre	1.152	0.122	0.913	1.390
Q3	Pre	1.167	0.124	0.924	1.409
Q4	Pre				
Q5	Pre	1.100	0.114	0.876	1.324
Q1	Post	1.278	0.124	1.034	1.521
Q2	Post	1.030	0.125	0.784	1.275
Q3	Post	0.956	0.131	0.700	1.211
Q4	Post	0.804	0.117	0.574	1.033
Q5	Post	1.189	0.116	0.962	1.415

Note. The results are averaged over the levels of *group*.

3.1.2.2.4 Effects on L4

Regarding the perception of L4 stimuli, the LMMs also found a significant interaction between *time* and *quality* (F(4, 266.69)= 2.867, p= 0.024). However, no main effects were obtained for *group, time*, or *quality.* The relationship between *time* and *quality* can be seen in more detail in Figure 6.9, where the *beat* identifications for Q1 are higher in the post-test than in the pre-test. Conversely, Q5 was more often identified as *bit* in the post-test. Table 4.9.2 shows that pre-test qualities were more often perceived half the time as *bit* than post-test, except for endpoint Q5 (pre-test, EMM = 1.237; post-test, EMM = 0.933).

Table 4.9.1. Results of LMMs on quality, group, and time effects on L4

Effect	df	F	p
Group	2, 50.74	0.637	0.533
Time	1, 41.24	1.191	0.281
Quality	4, 62.63	1.682	0.165
Group ＊ Time	2, 42.57	0.423	0.658
Group ＊ Quality	8, 65.46	0.709	0.682
Time ＊ Quality	4, 266.69	2.867	0.024
Group ＊ Time ＊ Quality	7, 266.69	0.577	0.774

Figure 6.9. Plot of vowel identification as a function of time and quality

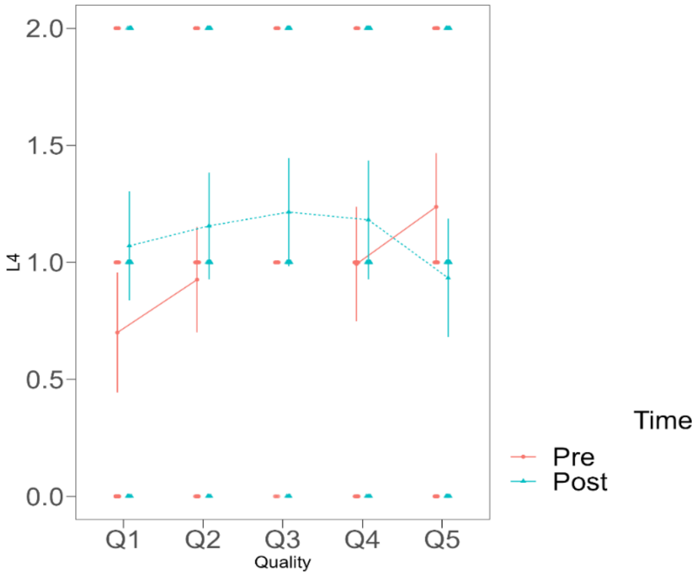


Table 4.9.2. Estimated Marginal Means for vowel identification across qualities

				95% CI	
Quality	**Time**	**Estimate**	**SE**	**Lower**	**Upper**
Q1	Pre	0.700	0.129	0.448	0.952
Q2	Pre	0.926	0.113	0.704	1.148
Q3	Pre				
Q4	Pre	0.993	0.123	0.752	1.234
Q5	Pre	1.237	0.115	1.011	1.463
Q1	Post	1.070	0.117	0.840	1.300
Q2	Post	1.156	0.115	0.931	1.380
Q3	Post	1.215	0.116	0.987	1.442
Q4	Post	1.181	0.127	0.932	1.431
Q5	Post	0.933	0.127	0.685	1.182

Note. The results are averaged over the levels of *group*.

3.1.2.2.5 Effects on L5

The LMMs for L5 stimuli returned no statistically significant effect for *group* (p = .275), *time* (p = .389), or *quality* (p = .229). Figure 6.10 illustrates the vowel identifications in the pre-test and post-test. The post-test *beat* identifications were consistently higher in the post-test than in the pre-test across stimuli that. Table 4.10.2 contains the EMMs of pre-test and post-test vowel identifications. The post-test EMMs seem to mirror the progression of the pre-test's EMMs in that they progressively climb as they approach the quality endpoint Q5; however, the post-test received higher overall *beat* identifications than the pre-test for the same stimuli (e.g., pre-test, Q5L5, EMM = 1.030; post-test, EMM = 1.122).

Figure 6.10 Plot of vowel identification as a function of time and quality

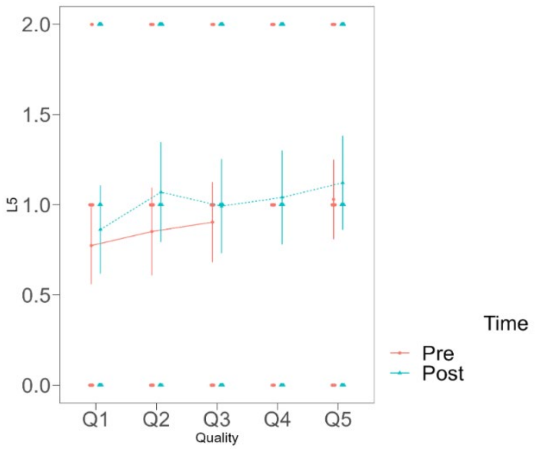


Table 4.10.2. Estimated Marginal Means for vowel identification across qualities

				95% CI	
Quality	**Time**	**Estimate**	**SE**	**Lower**	**Upper**
Q1	Pre	0.774	0.108	0.563	0.985
Q2	Pre	0.852	0.122	0.613	1.091
Q3	Pre	0.904	0.112	0.685	1.122
Q4	Pre				
Q5	Pre	1.030	0.111	0.812	1.247
Q1	Post	0.863	0.122	0.623	1.103
Q2	Post	1.070	0.139	0.799	1.342
Q3	Post	0.993	0.131	0.737	1.249
Q4	Post	1.041	0.130	0.786	1.295
Q5	Post	1.122	0.130	0.867	1.378

Note. The results are averaged over the levels of *group*.

3.1.2.3 Estimated marginal means analysis across group and time

The previous section reported the findings of the Linear Mixed Models including multiple significant main effects of variables across stimuli perceptions. Thus, it was deemed that an analysis of estimated marginal means of each stimulus across *time* and *group* by way of a post-hoc test would be worth investigating to better assess the differences of these variables.

Table 4.11.1 presents the estimated marginal means of subjects' vowel identification in the pre-test, which demonstrates how their vowel identification varied with the shortening of vowel duration, as indicated by the lighter shade of red and higher frequency of decimals. Table 4.11.2 reflects a similar trend in the post-test, albeit to a slightly lesser extent, showing marginally more accurate identifications of vowel quality towards the end of the duration spectrum (*bit)*. However, the stimuli with the original spectral values of *beat* (Q1L1) and *bit* (Q5L1) were misidentified to a degree in the pre-test and post-test, as Q1L1 received more *bit* identifications (pre-test, EMM = 1.004; post-test, EMM = 1.167) than Q5L1 (pre-test, EMM = 1.374; post-test, EMM = 1.326).

Table 4.11.1. Estimated Marginal Means of all groups in pre-test

		Duration				
		long				short
Quality	beat	1.004	1.052	0.863	0.7	0.774
		1.189		1.152	0.926	0.852
		1.293	1.374	1.167		0.904
		1.326	1.356		0.993	
	bit	1.374	1.319	1.1	1.237	1.03

beat 2

bit 0

Note. The results are averaged over the levels of *group*.

Table 4.11.2. Estimated Marginal Means of all groups in post-test

Duration

Quality					
beat	1.167	1.144	1.278	1.07	0.863
	1.185	1.341	1.03	1.156	1.07
	1.044	1.193	0.956	1.215	0.993
	1.233	1.007	0.804	1.181	1.041
bit	1.326	1.133	1.189	0.933	1.122

beat 2

bit 0

Note. The results are averaged over the levels of *group*.

A Linear Mixed Models was run on the estimated marginal means to assess the significance of *time* and *group*. The results show that some stimuli identifications were significantly influenced by *time*: Q1L3 (F(2, 40)= 7.719, p= .008) p = .008), Q1L4 (F(1, 40)= 4.854, p= .033), Q4L2 (F(1, 40)= 5.358, p = 0.026), indicating a significant change in vowel perception between the pre-test and the post-test. Tables 4.11.1–4.11.2 show that Q1L3 and Q1L4 received higher *beat* identifications in the post-test, thus more subjects correctly identified them as *beat* in the post-test, whereas Q4L2 (a spectrally ambiguous stimulus closer to *bit*) received significantly fewer *beat* identifications in the post-test (EMM = -0.349).

Furthermore, a significant interaction between *time* and *group* (F(2, 80)= 3.503) p= .035) was found on the perception of Q1L5. Figure 6.11 illustrates this interaction: while the QT and IT groups correctly perceived Q1L5 ("short" *beat*) more often in the post-test, the LT group identified Q1L5 as *bit* more often in the post-test. Furthermore, in regard to the estimated marginal means of Q1L4 identification, LT improved the least (EMM = + 0.111) in correctly identifying the stimulus as *beat* compared to QT (EMM = +0.400) and IT (EMM = +0.600). On the other hand, Q1L3 was more accurately perceived by LT (EMM = +0.611) and IT (EMM = +0.733) QT (EMM = -0.100).

The identification of Q3L1 was also found to be significantly affected by *time* ($F(2, 40)= 8.237$) $p= 0.001$). Figure 6.12 illustrates this main effect of *time* on identification of Q3L1. While the three groups identified the stimulus as *beat* less often in the post-test, the IT's *beat* identifications were higher in pre-test (EMM = 1.600) and post-test (EMM = 1.533) than those of the QT (pre-test, EMM = 1; post-test, EMM = 0.600) and LT (pre-test, EMM = 1.278; post-test, EMM = 1).

Figure 6.11. Plot of identifications of Q1L5 as a function of time and group

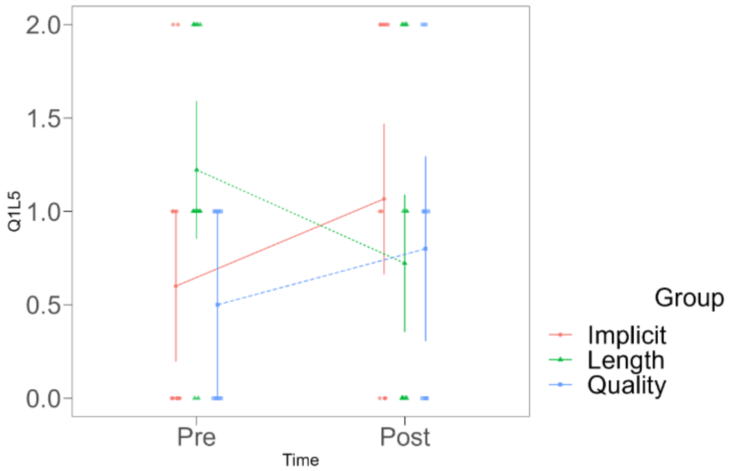
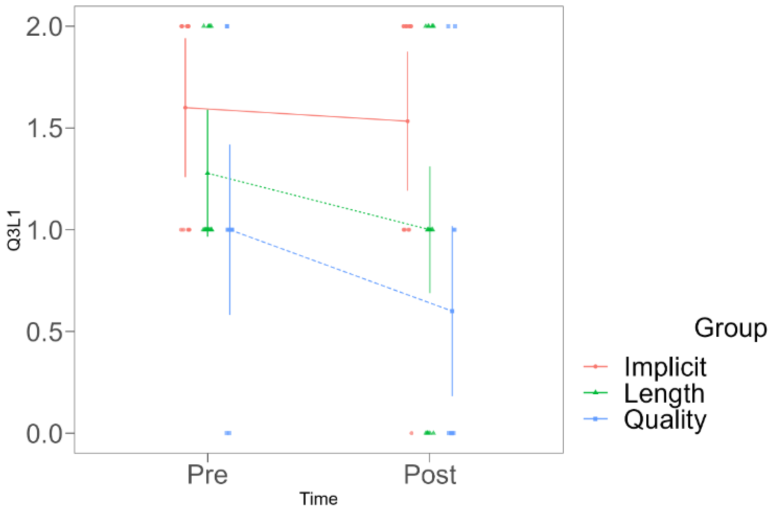


Figure 6.12. Plot of vowel identification of Q3L1 as a function of time and group



3.1.2.4 Differential analysis of perception scores across endpoints

In order to better assess the effect of duration and spectral manipulations, the differential analysis of vowel identifications as a function of length (L1, L5) and quality endpoints (Q1, Q5) was a carried out.

For the analysis of length differences, the number of vowel identifications (from 0 to 2) for long stimuli L1 were subtracted from their short counterparts L5 for each quality. Likewise, the vowel identifications (from 0 to 2) of Q1 were subtracted from their counterparts Q5. A Linear Mixed Models was used to test the data for main effects of *time, group, length difference*, and *quality difference.* However, no significant main effect was found (Tables 4.13.1 and 4.13.2). Therefore, it cannot be inferred that length differences and quality differences at the endpoints affected the perception of vowel identifications from subjects.

Table 4.13.1. Results of LMMs on group, time, length differences

Effect	df	F	p
Group	2, 40.55	2.405	0.103
Time	1, 40.70	0.872	0.356
Length difference	4, 84.00	0.065	0.992
Group ✻ Time	2, 41.44	1.756	0.185
Group ✻ Length difference	8, 88.70	1.835	0.081
Time ✻ Length difference	4, 261.70	0.631	0.641
Group ✻ Time ✻ Length difference	7, 261.70	0.284	0.960

Table 4.13.2. Results of LMMs on group, time, length differences

Effect	df	F	p
Group	2, 40.02	1.118	0.337
Time	1, 40.02	2.368	0.132
Quality difference	4, 75.29	0.664	0.619
Group ✻ Time	2, 40.02	0.480	0.622
Group ✻ Quality difference	8, 75.29	0.409	0.912
Time ✻ Quality difference	4, 276.06	1.815	0.126
Group ✻ Time ✻ Quality difference	8, 276.06	0.436	0.899

3.2. Production experiment

The analysis of the production data sought to shed light on RQ2: "To what extent do the different approaches to pronunciation help the participants improve their vowel perception and production?". To this end, three analyses were carried out to ascertain the level of improvement that was observable in subjects' tense-lax vowel productions: 1) effect of training on F1, F2 and duration, 2) relative F1-F2 distance, and 3) relative tense-lax distance. Importantly, all the analyses introduced the variable *native-likeness* (native speakers' F1, F2 and duration measurements) to ascertain the subjects' proximation to native-like vowel productions as measured in F1, F2, and duration. Furthermore, initial tests included the variable *model* (the speech stimuli that was presented to subjects after they produced the tokens and before repeating them). However, to simplify the analysis and complexity of the data set, this variable was excluded, since no main effect was found in any of the three tests with Linear Mixed Models.

3.2.1 Vowel plots

Figures 7.1.1–7.2.6 are vowel plots of productions of high front tense-lax and high back tense-lax vowels at pre-test and post-test from IT, QT, and LT. For each vowel plot, the F1, F2, and F3 of each participant were used to plot each vowel production using Wendy's (2021) web-based vowel plotting tool.

Regarding high back vowel productions in the pre-test (Figures 7.1.1, 7.1.3, and 7.1.5), tense (green) and lax (red) vowels seem to concentrate on roughly the same area of the vowel space across the three groups' average productions. Additionally, the relative position of lax vowels in relation to their tense counterparts suggests that subjects did not produce native-like spectral contrasts. Specifically, average lax productions of all groups in the pre-test seem to exhibit higher F2 values since they appear to be higher in the vowel space than their tense counterparts. Native speakers, however, articulate high front lax vowel with lower F2 values and, thus, lower in the vowel space than high front tense.

On the other hand, post-test productions of high front tense-lax vowels seem to elicit an incipient effort in producing qualitatively contrastive vowels (Tables 7.1.2, 7.1.4, and 7.1.6). Most noticeably, IT and LT produced high front lax vowels with higher F1 and lower in F2 on average when compared to tense vowel productions, and so did QT albeit to a lesser extent. Consequently, the position that tense and lax vowel productions had in the pre-test reversed and adopted a more native-like placement (i.e., lax vowels pronounced lower and more retracted than tense counterparts).

In contrast, high back tense-lax vowels do not seem to separate to such a degree as seen in high front vowels. Figures 7.2.1, 7.2.3, and 7.2.5 seem to reflect a similar lack of vowel separation as seen in the high front vowels in pre-test. However, Figures 7.2.2, 7.2.4, and 7.2.6 show a slight vowel separation and a more native-like orientation, in that lax vowels appear to be higher in F2 than tense counterparts.

Figure 7.1.1 Vowel plots of high front tense-lax vowels from IT in pre-test

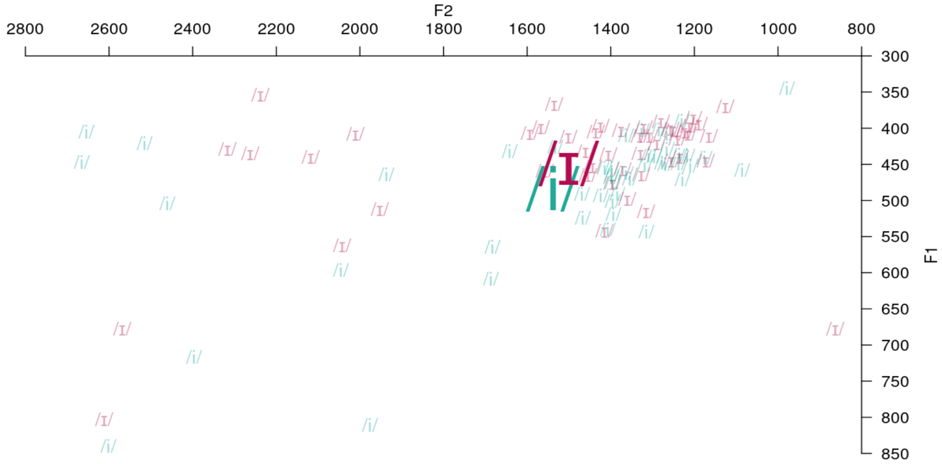


Figure 7.1.2 Vowel plots of high front tense-lax vowels from IT in post-test

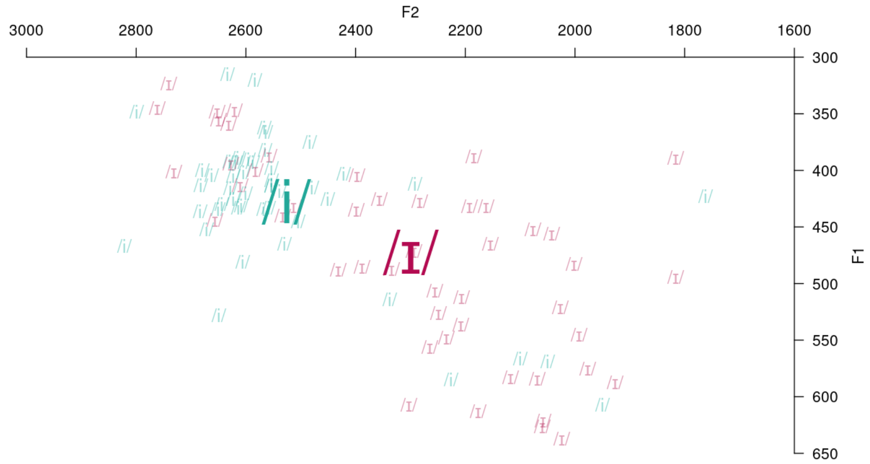


Figure 7.1.3 Vowel plots of high front tense-lax vowels from QT in pre-test

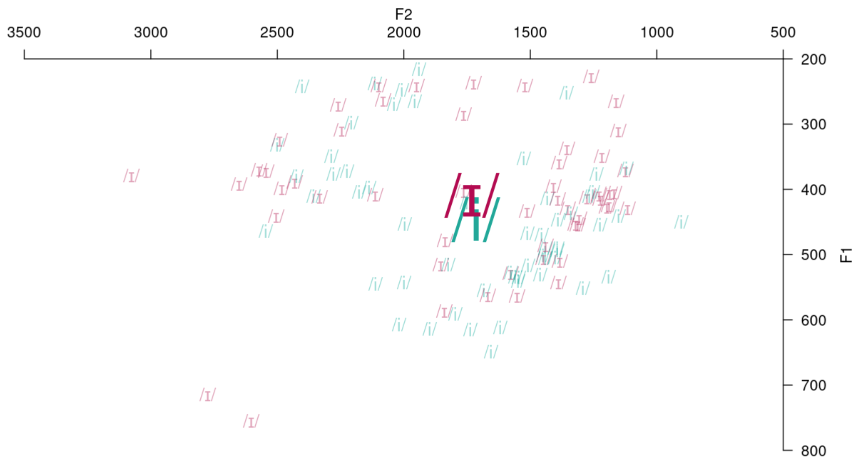


Figure 7.1.4 Vowel plots of high front tense-lax vowels from QT in post-test

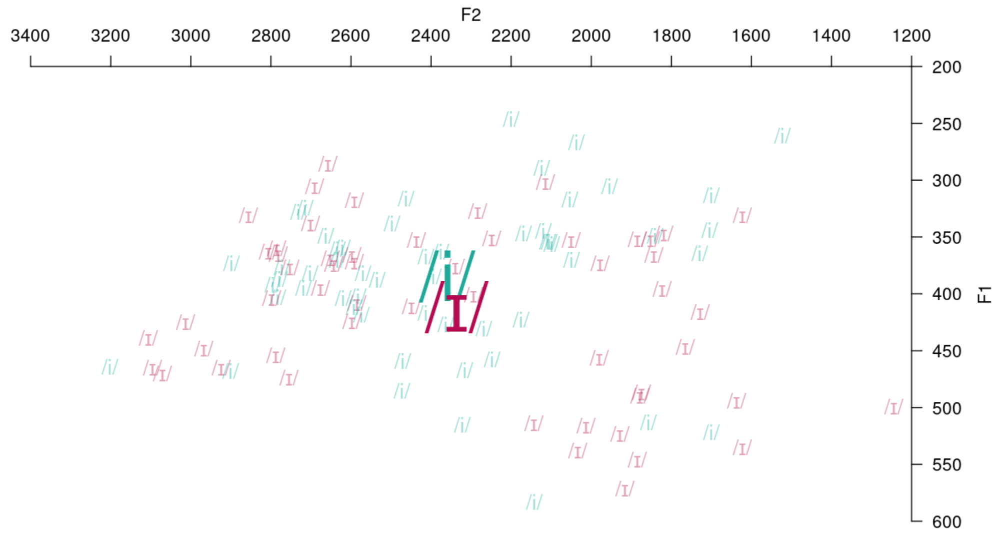


Figure 7.1.5 Vowel plots of high front tense-lax vowels from LT in pre-test

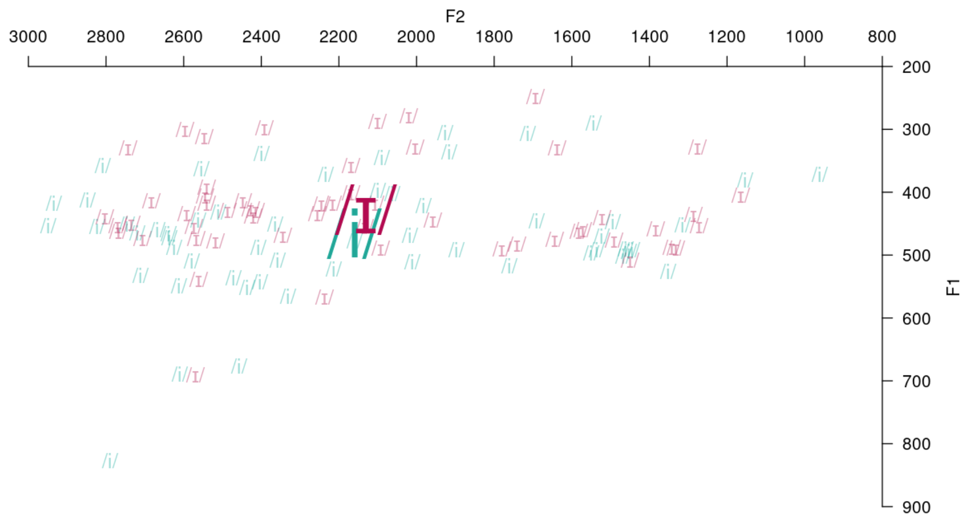


Figure 7.1.6 Vowel plots of high front tense-lax vowels from LT in post-test

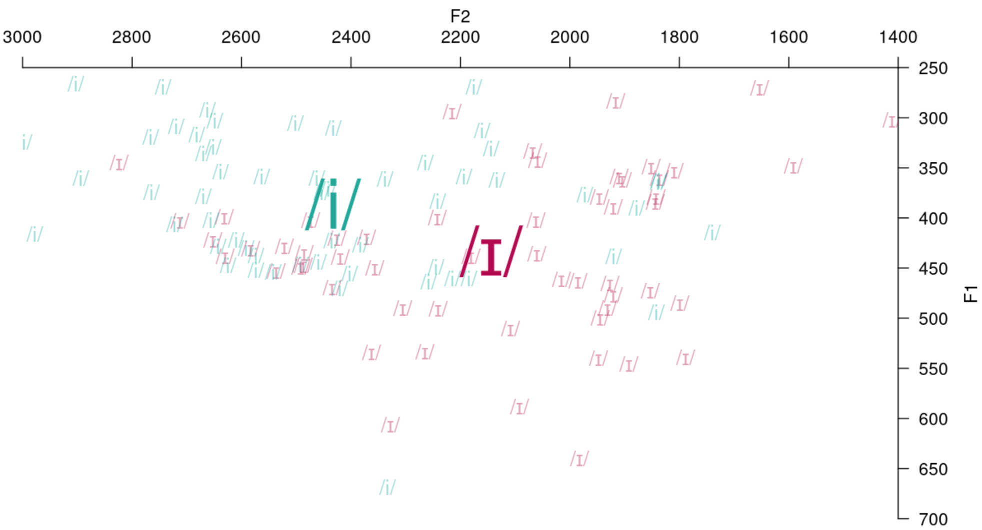


Figure 7.2.1 Vowel plots of high back tense-lax vowels from LT in pre-test

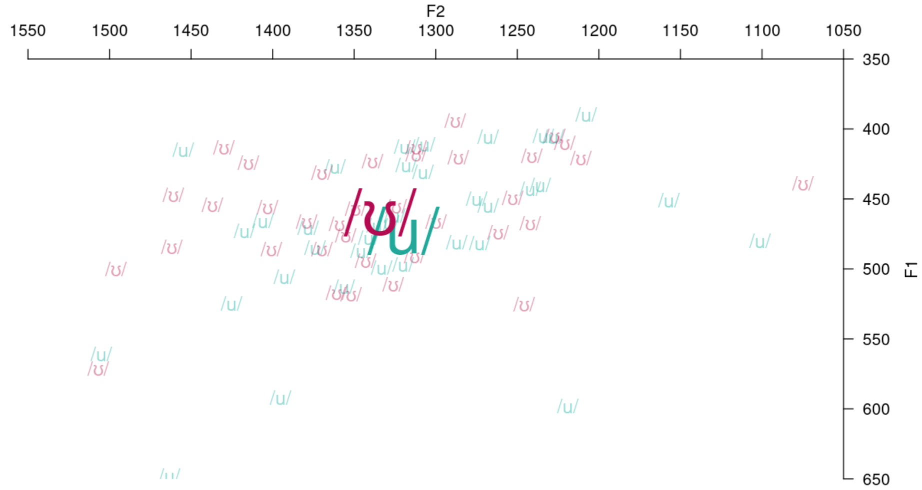


Figure 7.2.2 Vowel plots of high back tense-lax vowels from LT in post-test

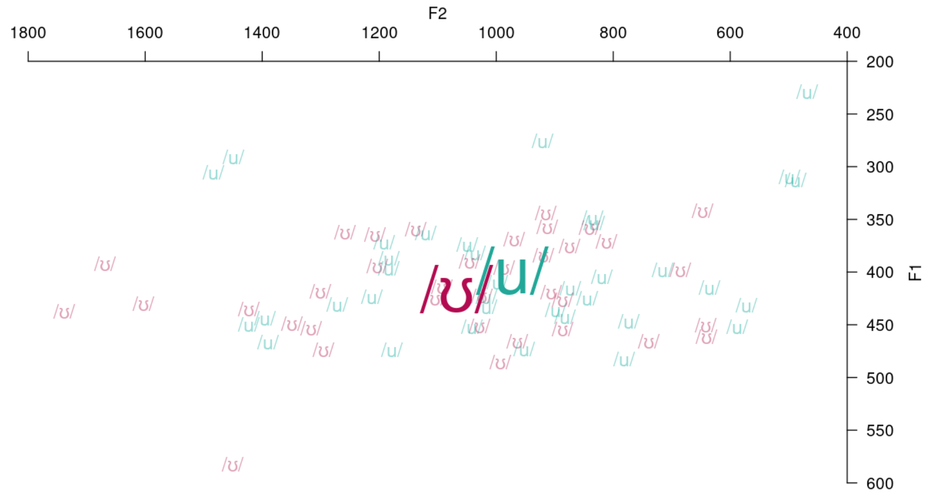


Figure 7.2.3 Vowel plots of high back tense-lax vowels from QT in pre-test

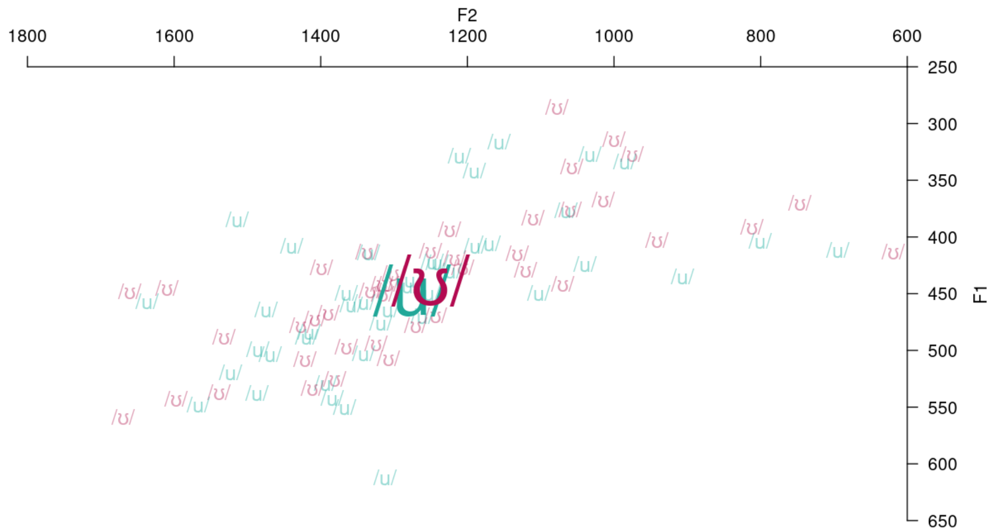


Figure 7.2.4 Vowel plots of high back tense-lax vowels from QT in post-test

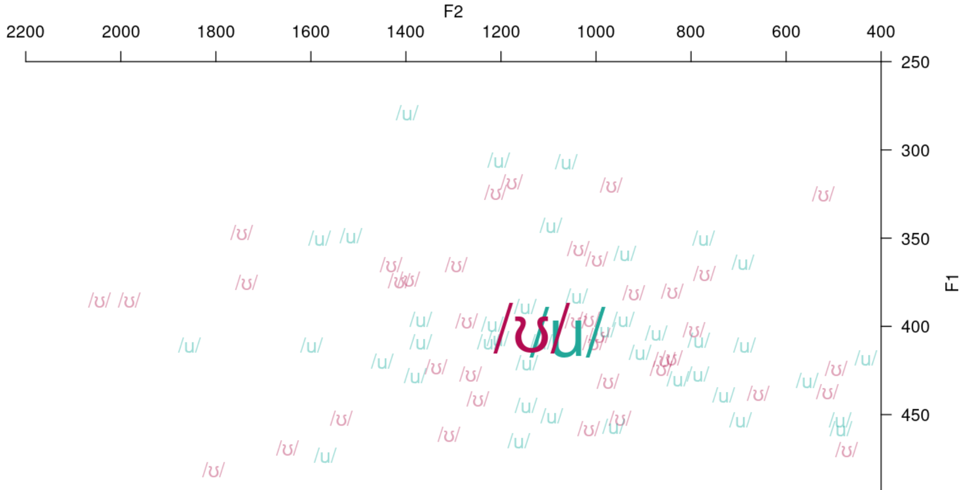


Figure 7.2.5 Vowel plots of high back tense-lax vowels from LT in pre-test

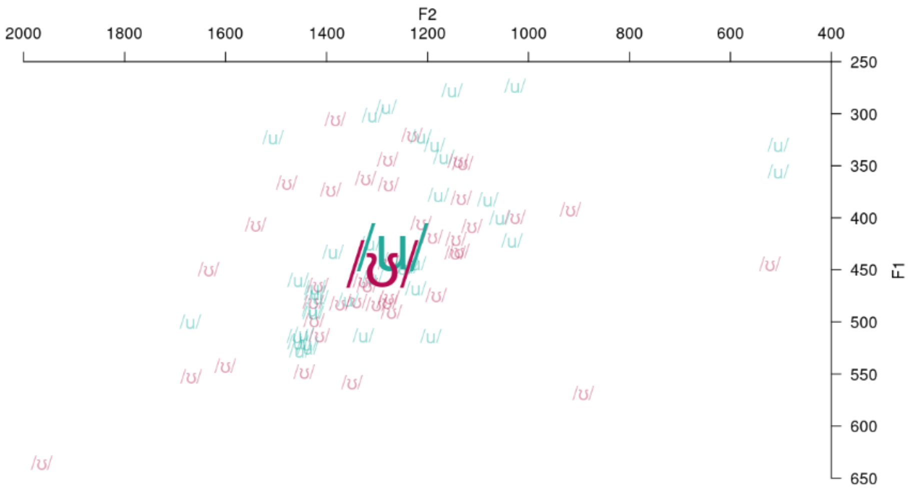
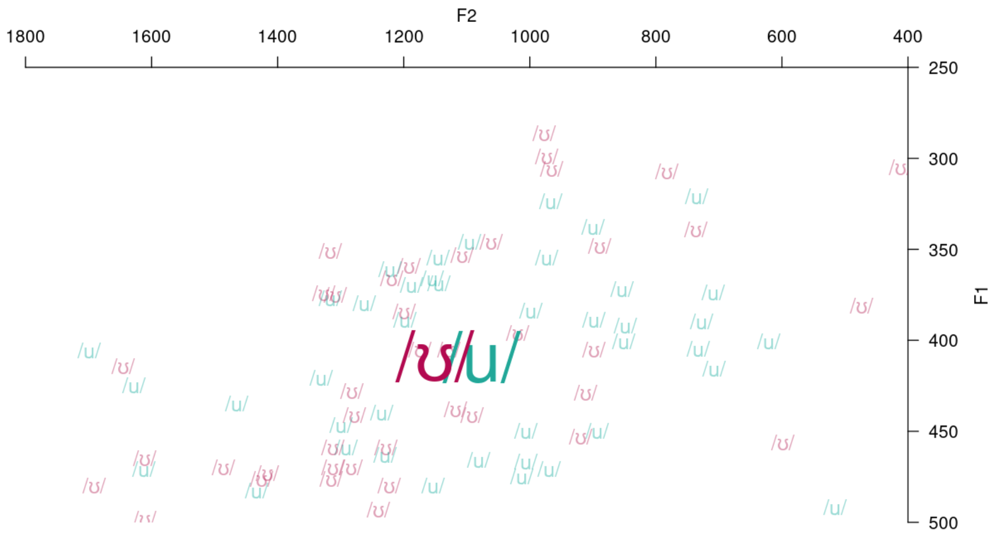


Figure 7.2.6 Vowel plots of high back tense-lax vowels from LT in post-test



3.2.2 Effect of training on F1, F2, and duration

A Linear Mixed Models aimed to test the effect of *time* and *group* variables on two acoustic dimensions in subjects' tense-lax productions; namely, spectra (F1 and F2 measured in Hz) and duration. Additionally, a third variable, *native-likeness*, was introduced by way of a benchmark comparison to assess participants' proximity to native speakers' acoustic dimensions of tense-lax articulations. In this regard, *native-likeness* effect of $p > .005$ was interpreted as positive closeness to native speakers' acoustic values of F1, F2, and duration.

3.2.2.1 High front tense

The LMM returned significant effects for *time* ($p< .001$) and found a significant interaction between *time* and *native-likeness* ($p< .001$). Figure 7.3.1 illustrates the effect of time and the distance to *native-likeness*. While the estimated marginal mean of F1 values from the non-native speakers (NNS) was 452 in the pre-test, that of the native speakers (NS) was 465.3. However as can be seen in the figure all three groups separate significantly from native F1 values in the post-test (NNS, EMM = 399.8). This indicates that subjects' high front tense vowels were pronounced with lower F1 (higher in the vowel space). On the other hand, the LMMs found significant effects of all the variables and their interactions on the F2 of productions: *group* ($p = .044$), *time* ($p < .001$), *native-likeness* ($p < .001$), *group*time* ($p = .004$), *group*native-likeness* ($p = .044$), *time*native-likeness* ($p < .001$), and *group*time*native-likeness* ($p = .004$).

Figure 7.3.2 displays the effects of these variables: all groups' F2 approximated the native F2 value (3081) in the post-test; however, IT produced the most native-like F2

(EMM = 2526.52) in the post-test and LT exhibited the least change in F2. Therefore, all subjects displayed higher levels of frontness (F2) for /i/ in the post-test.

Regarding duration, only *native-likeness* ($p < .001$) was found to be significant. Figure 7.3.3 shows that subjects vowel durations (pre-test, EMM = 162; post-test, EMM = 174) was consistently far from native's duration values (297 ms) in pronouncing high front tense vowels in the same context as the tokens used in the production experiment.

Figure 7.3.1 Plot: F1 from IT, LT, and QT high front tense tokens across time

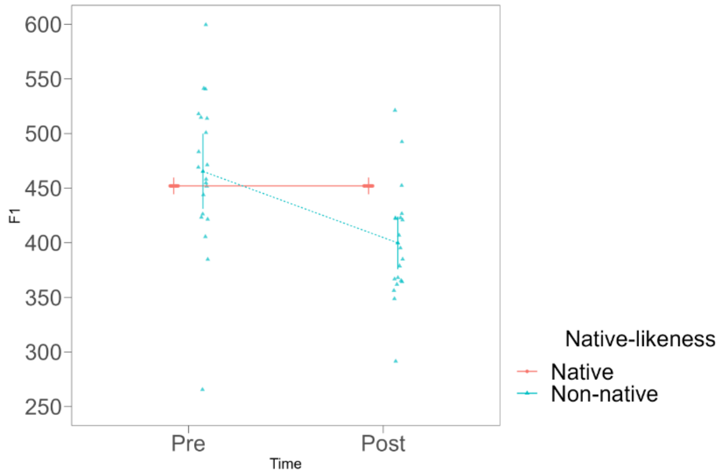


Figure 7.3.2 Plot: F2 from IT, LT, and QT high front tense tokens across time

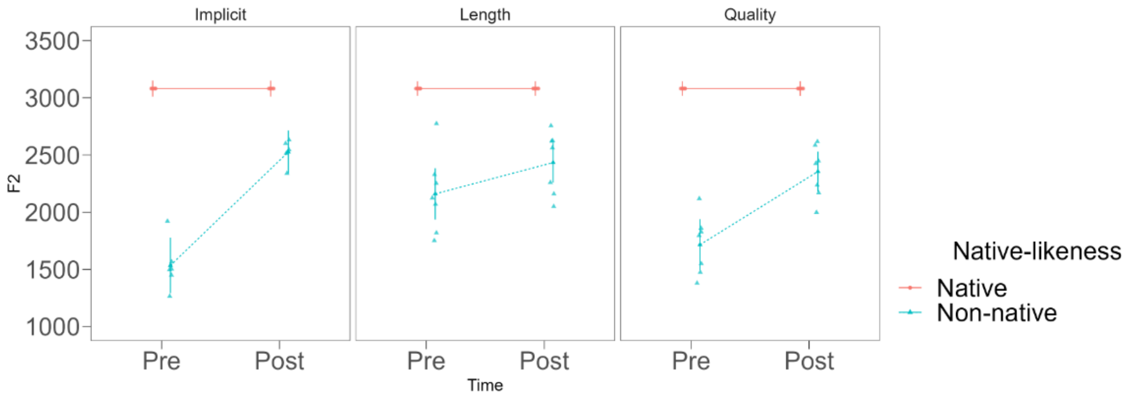
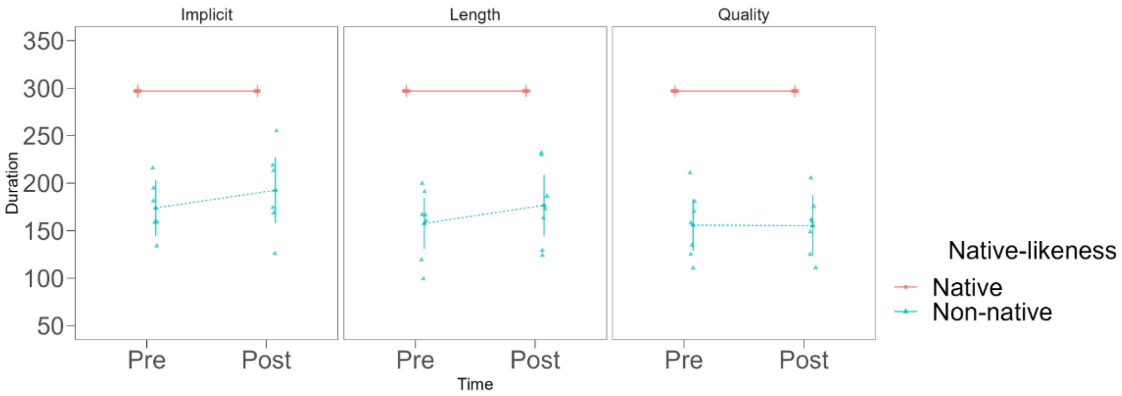


Figure 7.3.3 Plot: Duration from IT, LT, and QT high front tense tokens across time



3.2.2.2 High front lax

The LMMs only found *native-likeness* to be significant (p < .001) and Figure 7.3.4 shows that this significance stems from the consistent distance between native F1 (511) and subjects' F1 (pre-test, EMM = 429; post-test, EMM = 439). On the other hand, there were significant effects on F2 from *group* (p = .092), *time* (p < .01), *native-likeness* (p < .001), *group*time* (p = .036), *group*native-likeness* (p < .001), *group*time*native-likeness* (p = .036).

Figure 7.3.5 shows that the QT and IT groups approximated native F2 values (2552) in the post-test; however, LQ produced the F2 values (EMM = 2347) farthest from native ones. Regarding duration, the significant variables found were *time* (p =.039), *native-likeness* (p < .001), and the interaction *time*native-likeness* (p = .039). Figure 7.3.6 reveals a minimal change in duration after the intervention, but subject still did not produce native-like duration values (248 ms)

Figure 7.3.4 Plot: F1 from IT, LT, and QT high front lax tokens across time

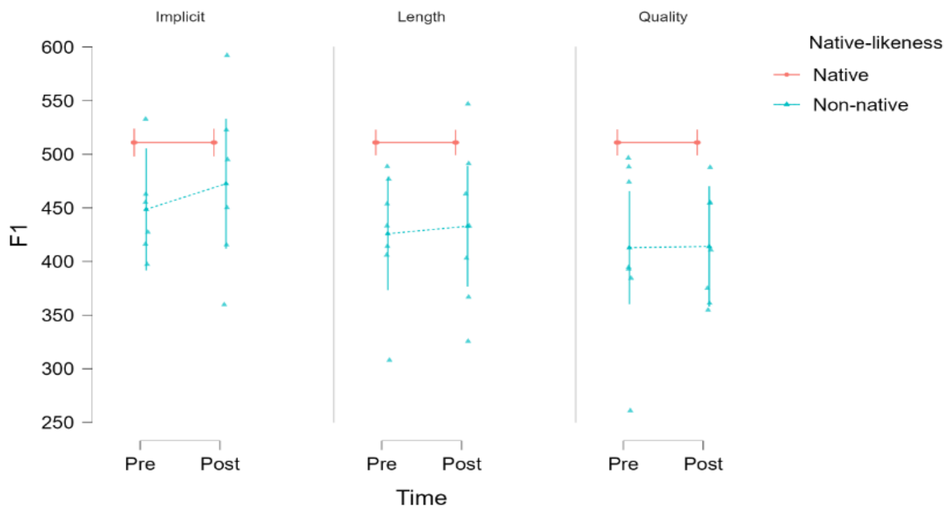


Figure 7.3.5 Plot: F2 from IT, LT, and QT high front lax tokens across time

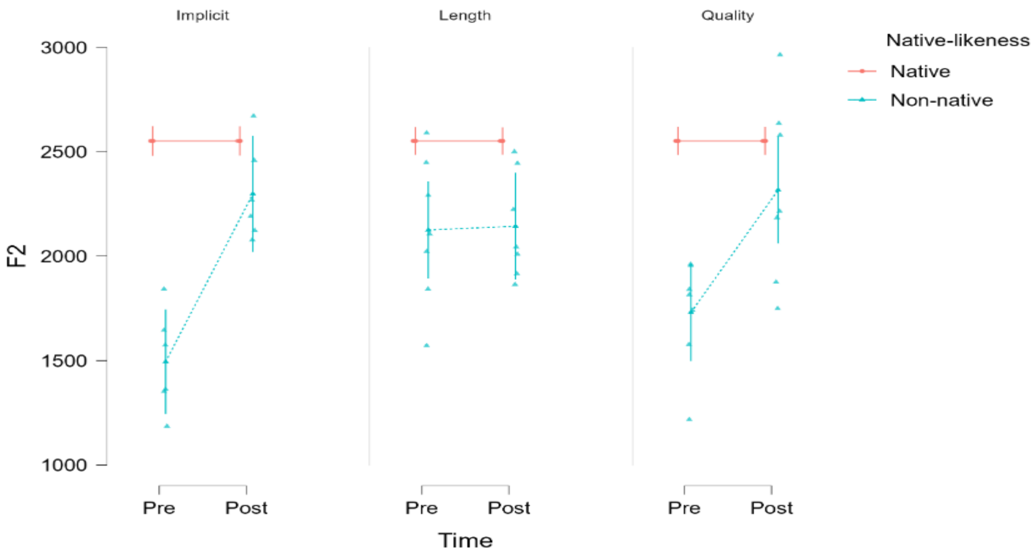
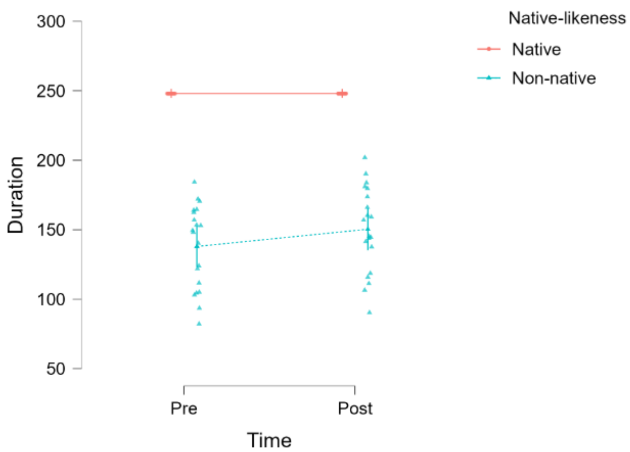


Figure 7.3.6 *Plot: Duration from IT, LT, and QT high front tense tokens across time*



3.2.2.3 High back tense

The LMMs found significant effect on F1 from *time* ($p < .001$), *native-likeness* ($p < .001$), and *time*native-likeness* ($p < .001$). Figure 7.3.7 illustrates how high back tense vowels deviated form native F1 values in the post-test. This change was spread across the three groups. Moreover, *time* ($p < .001$), *native-likeness* ($p < .001$), and *time*native-likeness* ($p < .001$) were found to be significant in the F2 of vowel productions.

According to Figure 7.3.9, the change in F2 in the post-test manifested as another deviation from native F2 values. With respect to duration, *native-likeness* ($p < .001$), *group*time,* and *group*time*native-likeness* were found to be significant effects on duration values. As can be seen in Figure 7.3.9, LT and IT produced durations for high back tense vowels (LT, IT, pre-test, EMM = 162) that were closer to native durations (278 ms).

Figure 7.3.7 Plot: F1 from IT, LT, and QT high back tense tokens across time

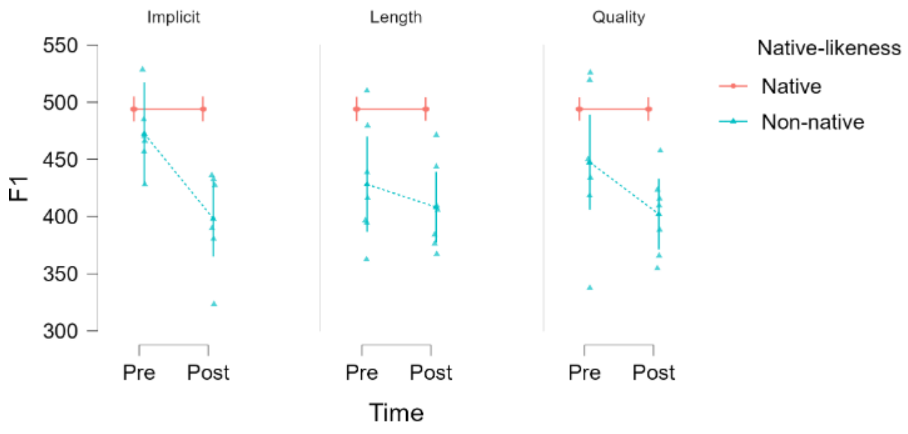


Figure 7.3.8 Plot: F2 from IT, LT, and QT high back tense tokens across time

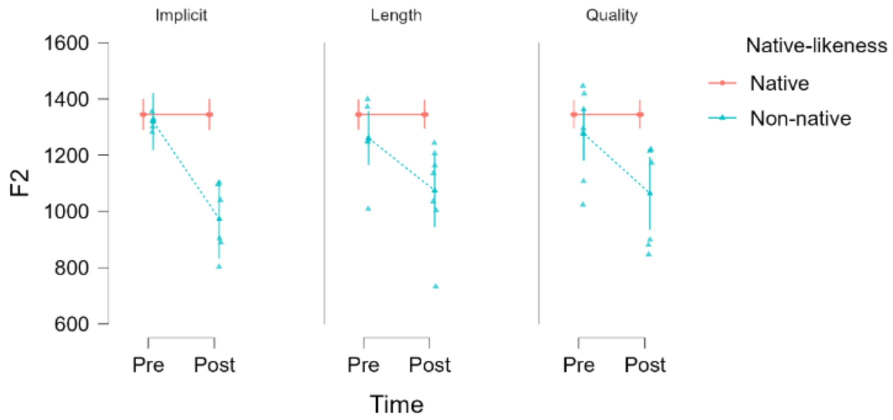
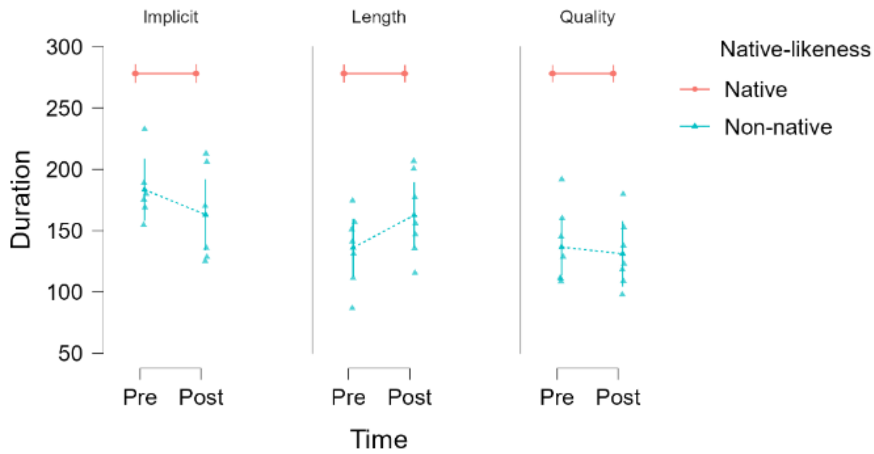


Figure 7.3.9 Plot: Duration from IT, LT, and QT high back tense tokens across time



3.2.2.4 High back lax

The LMMs found *time* ($p < .001$), *native-likeness* ($p < .001$), and *time*native-likeness* ($p < .001$) significant effects of F1 values for high back lax productions. The nature of these effects can be seen in Figure 7.3.10, which depicts a significant deviation from native F1 values in a downward trend, thus, pronouncing high back lax vowels higher in the vowel space. Likewise, *time* ($p = .013$), *native-likeness* ($p < .001$), and *time*native-likeness* ($p = .013$) were found to be significant regarding F2 values. Figure 7.3.11 represents the same downward deviation as seen in Figure 7.3.10, where high back lax productions had lower F2, thus pronounced more back. With respect to duration, only *native-likeness* was significant ($p < .001$; Figure 7.3.12 shows a stable duration values across groups and time.

Figure 7.3.10 Plot: F1 from IT, LT, and QT high back lax tokens across time

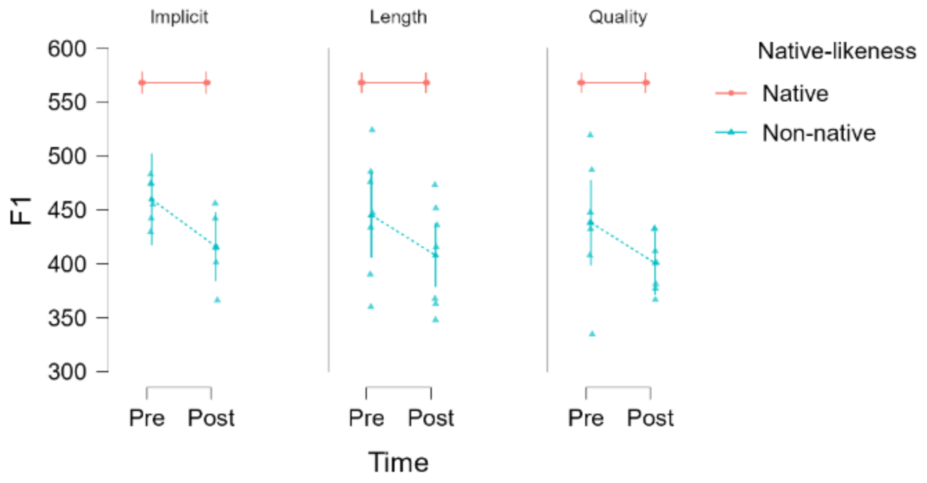


Figure 7.3.11 Plot: F2 from IT, LT, and QT high back lax tokens across time

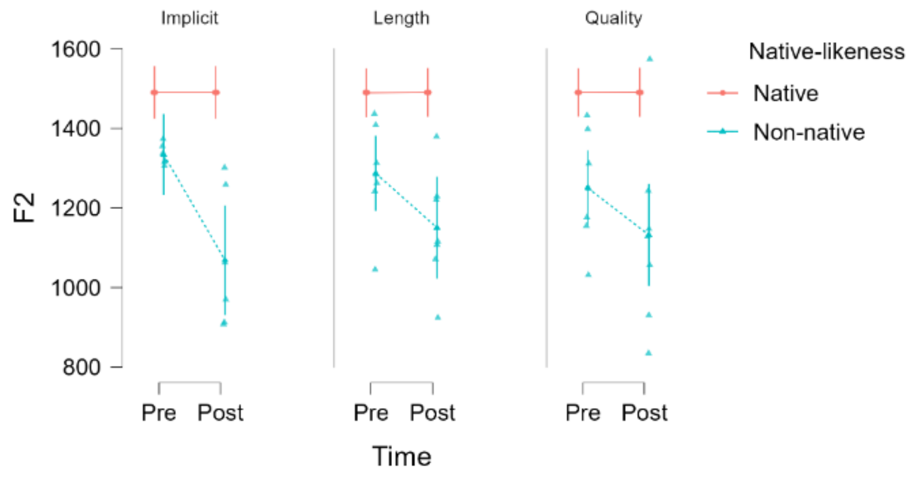
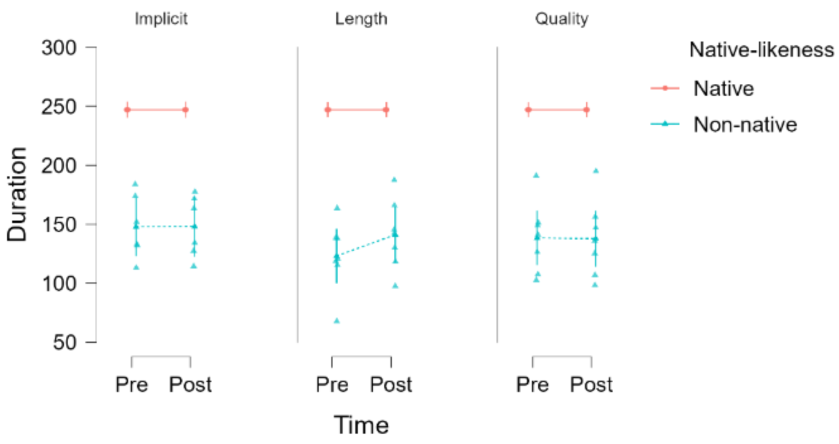


Figure 7.3.12 Plot: Duration from IT, LT, and QT high back lax tokens across time



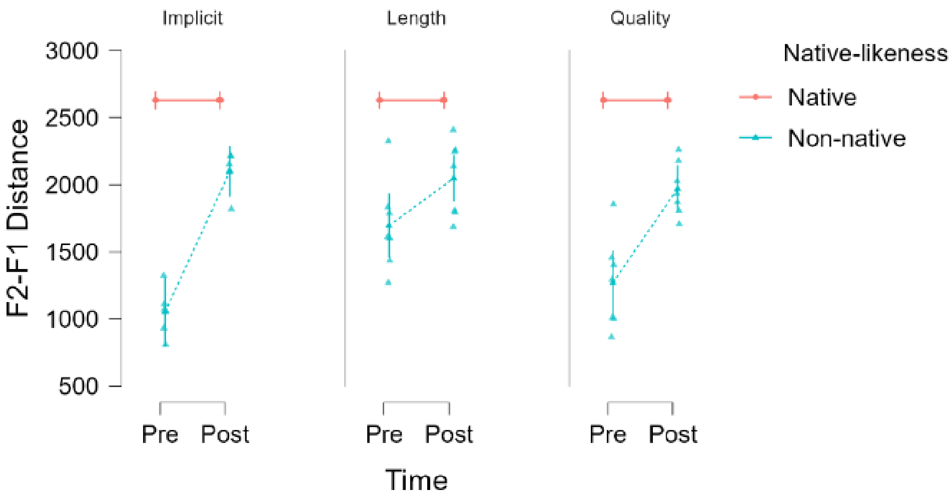
3.2.3 Relative F1-F2 distance

Since absolute F1, F2, and distance values were found to deviate from those of native speakers, despite the apparent improvement observed in the vowel plots, another measure was deemed appropriate to assess production improvement. The F1 of each vowel was subtracted from their F2, as an indication of relative F1-F2 distance, and then compared to native values.

3.2.3.1 Relative F1-F2 distance high front tense

The LMMs returned significant effects from all variables: group ($p = .043$), time ($p < .001$), native-likeness ($p < .001$), group*time ($p < .008$), Group*Native-likeness ($p < .043$), Time*Native-likeness ($p<.001$), Group*Time*Native-likeness ($p < .008$). Figure 7.4.1 shows an approximation to native F1-F2 distance after the pronunciation sessions. Subjects F1-F2 distance increased from the pre-test (EMM = 1337) to the post-test (EMM = 2038).

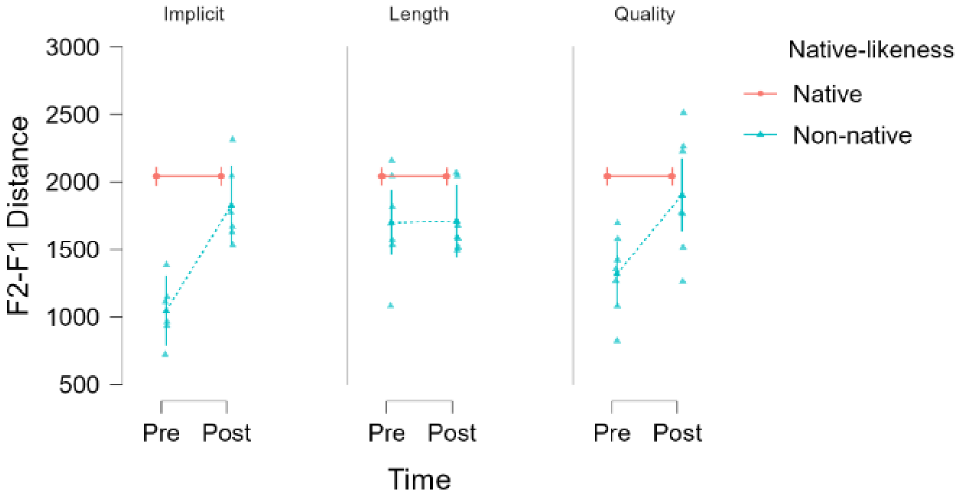
Figure 7.4.1 Plot: F1-F2 distance IT, LT, and QT high back tense tokens across time



3.2.3.2 Relative F1-F2 distance high front lax

Regarding high front lax vowels, the LMMs found significant effects from *time* (p = .002), *native-likeness* (p < .001), Time*Native-likeness" (p = .002), and near significant effect in *group*native-likeness* (p = .053). Figure 7.4.2 reveals that native's F1-F2 distance (2041) were approximated by those of IT and QT mainly, while LT's high front lax spectral area remained unchanged after the pronunciation sessions.

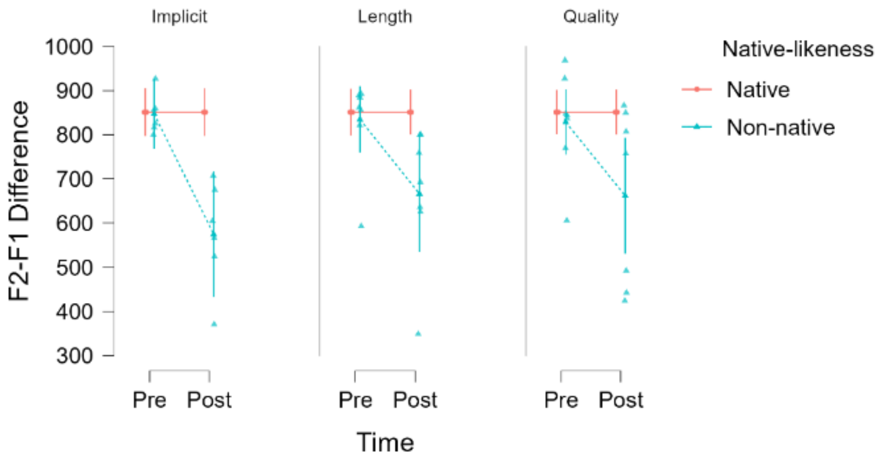
Figure 7.4.2 Plot: F1-F2 distance IT, LT, and QT high front lax tokens across time



3.2.3.3 Relative F1-F2 distance high back tense

The variables found significant in the LMMs were *time* (p < .001), *native-likeness* (p < .001), and *time*native-likeness* (p < .001). According to Figure 7.4.3, subjects' F1-F2 distances were significantly closer to those of natives than in the post-test. The F1-F2 in the post-test decreased, indicating a more concentrated, smaller spectral area in producing high back tense vowels.

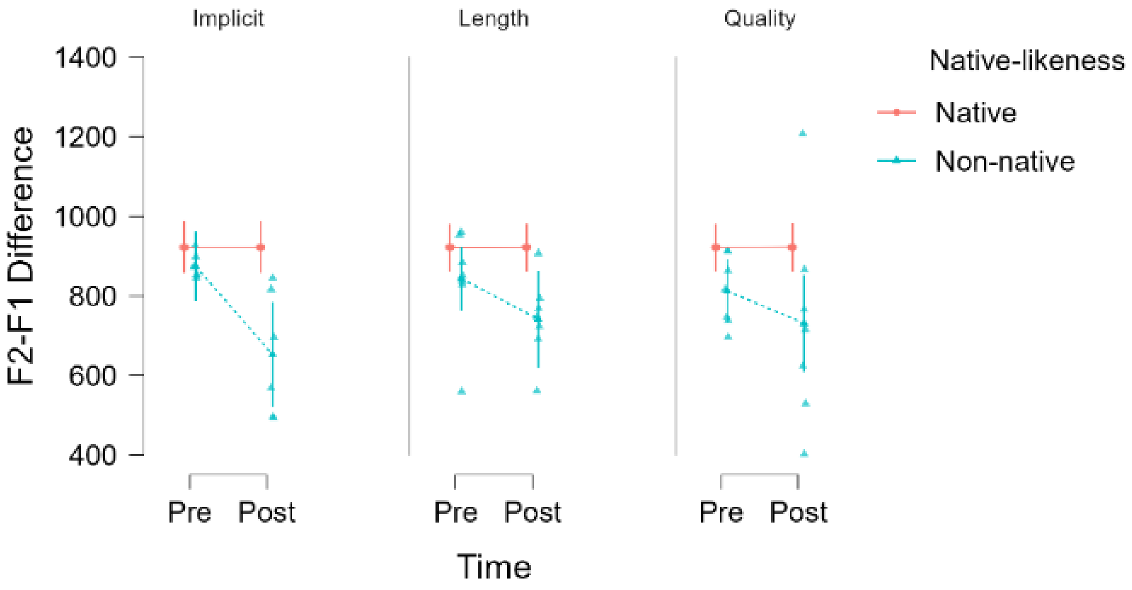
Figure 7.4.3 Plot: F1-F2 distance IT, LT, and QT high back lax tokens across time



3.2.3.3 Relative F1-F2 distance high back lax

The F1-F2 distance in high back lax productions was significantly affected by *time* (p = .032); *native-likeness* (p < .001) and *time*native-likeness* (p = .030) were also found significant. As can be seen in Figure 7.4.4, there is a downward trend in F1-F2 distances across the three groups, thus deviating from native values. This indicates that subjects' high back lax vowels were produced in a more smaller spectral area of the vowel space in the post-test.

Figure 7.4.4 Plot: F1-F2 distance IT, LT, and QT high back lax tokens across time



3.2.3.4 Relative tense-lax distance

Another measure that was considered insightful was the distance of acoustic dimensions between tense and lax counterparts. An LMMs was carried out on the effects of *time* and *group*.

3.2.3.4.1 Relative tense-lax distance F1 high front

Some of the significant variables found in the test on F1 were *time* (p < .001), *native-likeness* (p < .001), and *time*native-likeness* (p <.001). As Figure 7.5.1 reveals, the pronunciation training had an effect on subjects' distance between their tense and lax productions, thus approximating native tense and lax distances regarding vowel height. Moreover, the LMMs also revealed *native-likeness* and *time*native-likeness* to be significant variables. According to Figure 7.5.2 these variables can be understood to influence subjects in such a way that they produce more native-like tense-lax F2 contrasts.

Figure 7.5.1 Plot: High front tense-lax distance of F1 across group and time

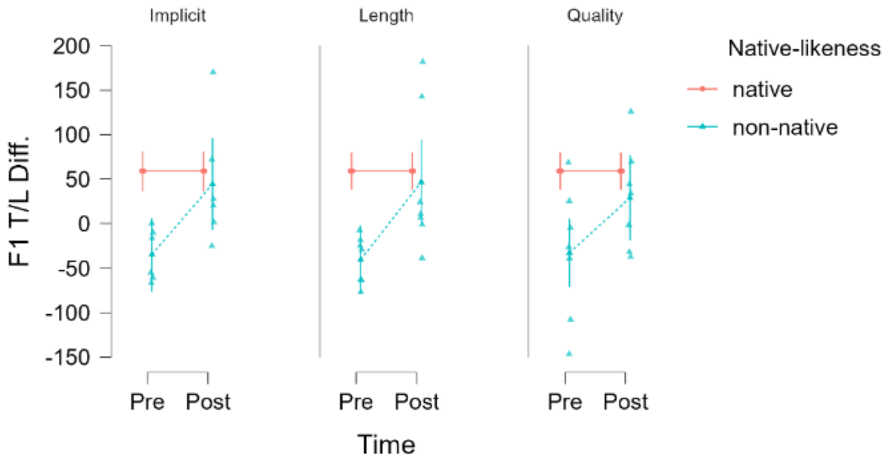
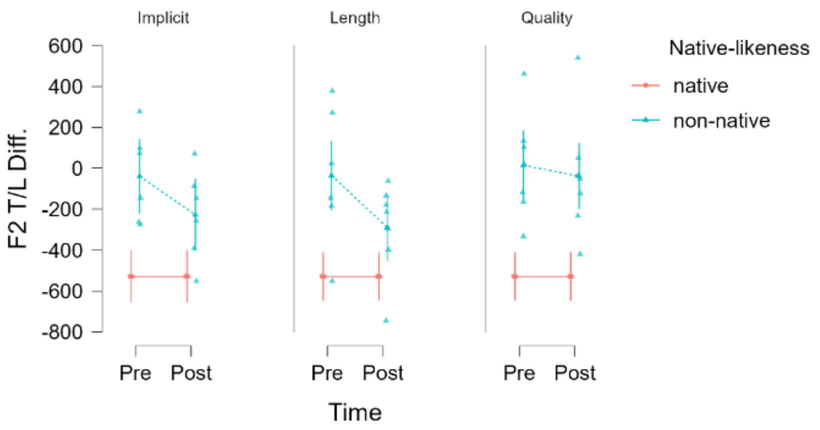
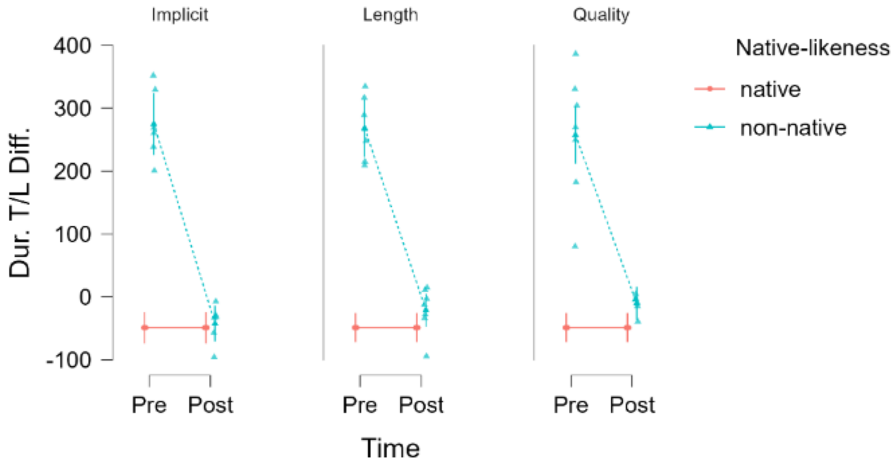


Figure 7.5.2 Plot: High front tense-lax distance of F2 across group and time



Moreover, the LMMs found *time* (p < .001), *native-likeness* (p < .001), and *time*native-likeness* (p < .001) significant variables. According to Figure 7.5.3 the duration measurements of subjects significantly approximated those of native speakers after the interventions.

Figure 7.5.3 Plot: High front tense-lax distance of duration across group and time



3.2.3.4.4 Relative tense-lax distance F1 high back

The only variables that were found significant were *native-likeness* (p < .001) and *group*time*native-likeness* (p = .013). Figure 7.5.4 shows that groups remain at distance from native values both in the pre-test and in the post-test. The significant variables found regarding F2 were *time* (p = .049), *native-likeness* (p < .001), and *time*native-likeness* (p = .003). Figure further illustrates the positive approximation of differences in F2 compared to native's distances after the pronunciation trainings. Finally, with regard to duration, *group* ($p < .001$), *native-likeness* ($p < .001$), *time*group* ($p = .041$) were found to be the significant effects. According to Figure 7.5.6, the change the IT and QT exceed native values, the LT approximated native duration values.

Figure 7.5.4 Plot: High back tense-lax distance of F1 across group and time

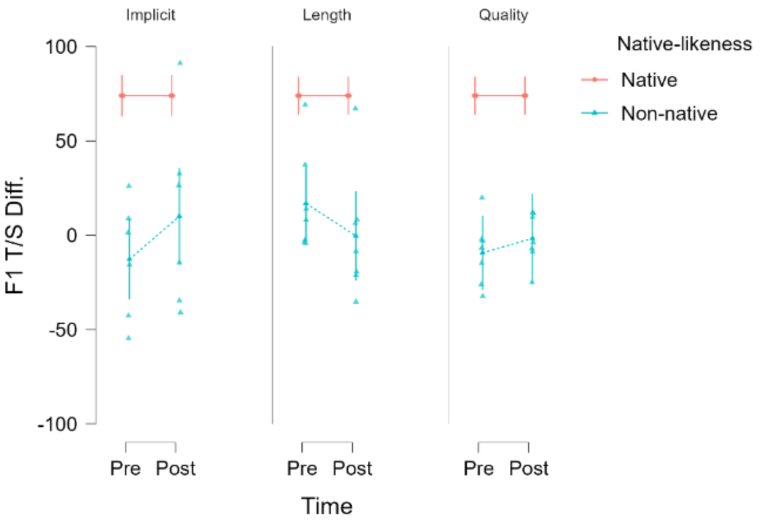


Figure 7.5.5 Plot: High back tense-lax distance of F2 across group and time

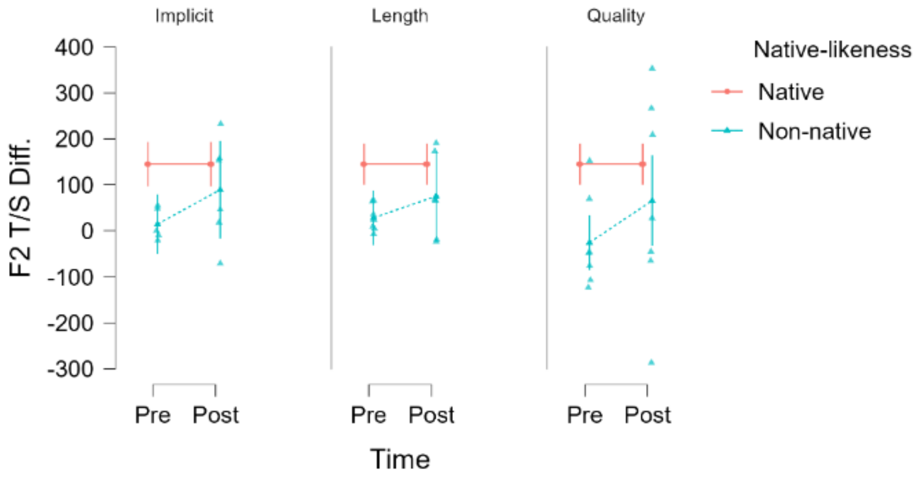
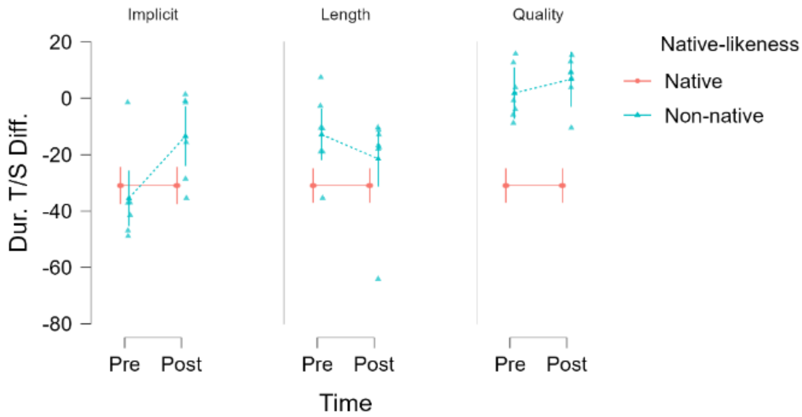


Figure 7.5.6 Plot: High back tense-lax distance of duration across group and time



Chapter 4 Discussion and conclusions

The analyses of *beat* identifications have provided substantial results that lend themselves to insights into how Spanish/Catalan speakers perceive non-native vowel pairs and the effect of short, but intensive pronunciation training.

4.1 Patterns in length and quality variations and vowel perception

Descriptive analyses seemed to show a pattern throughout the synthetic continuum, that of reliance on duration as a phonetic cue that subjects relied on to distinguish the tense-lax vowel pair. This trend was most noticeable in the IT group. However, statistical analysis revealed that only certain length variations significantly influenced how subjects perceived vowel sounds. This was demonstrated by the LMMs performed on the *beat* identification scores across lengths variations.

Regarding quality variations, another LMMs was run on the *beat* identification scores and only found a significant interaction between *time* and *quality*. This suggests that even though the effect of quality was not widespread in the pre-test or post-test, it did have an effect on the post-test to a lesser extent than was expected in the prediction to RQ2. Moreover, RQ3 cannot be supported, since QT subjects were not observed to be affected by vowel quality in the identification task after the intervention.

Moreover, even though quality was not found to be a significant variable, the descriptive analysis suggested that subjects consistently identified the stimuli endpoints Q1 and Q5 the other way round. That is, they demonstrated that 1) they were not aware of the spectral characteristics of each vowel and 2) they may have been confused by the unfamiliar spelling of "beat" and, in reading "bit" they assimilated that spelling to the pronunciation and spelling of Spanish /i/. In this regard, the effect of spelling may explain why Q5 was perceived as *beat* more often and Q1 as *bit*, since the one-to-one correspondence between spelling and pronunciation in Spanish may have primed inexperienced learners to associate the grapheme *"*i" in *bit* to the high front tense vowel which is more similar to the Spanish vowel than the lax counterpart.

Additionally, a further analysis afforded a deeper look into the influence of spectra and duration in subjects' vowel identification. Differential pre-test/post-test analysis indicated suggested intriguing insights into the shifts that subjects' perception underwent after the intervention. For example, the IT group exhibited increased *beat* identifications across stimulus variations of quality and length in the post-test, perceiving stimuli qualitatively closer to *beat* as the tense vowel. Conversely, the QT group demonstrated a decrease in *beat* identifications for certain stimuli, particularly those associated with qualitatively ambiguous sounds.

Therefore, RQ1 can be answered as follows: Spanish/Catalan speakers who do not have experience *may* rely on duration, especially for spectrally ambiguous stimuli.

4.2 Patterns of production improvement

The vowel productions of high front tense were shown to improve in different ways. Firstly, the F1 and F2 of tense and lax counterparts became further apart in the post-test, indicating an incipient recategorisation of non-native vowel sounds. Since, all subjects improved in this regard, RQ2 can be answered as follows: subjects of all groups improved the pronunciation of high front tense and lax vowels with regards to the relative distance between F1 and F2. Thus, typology of teaching pronunciation did not matter, since the three groups improved independently of training approach, even though the length group demonstrated higher degree of native-likeness as regards duration cues of tense-lax productions.

Where direct correspondence with native F1 and F2 values were not met, other measurements such as F1-F2 distance proved an improvement in vowel production in the post-test.

4.3 Conclusions

In this study, the perceptual cue weighting of duration and spectra were studied. Several key findings emerged from the analysis of vowel identification data, which suggest that Spanish/Catalan speakers *can* rely on duration. In the experiment, however, subjects relied more often on duration than on quality, as demonstrated in the statistical analysis of the experiment.

A pattern was observed in which the expected identification of *beat* was given to stimuli with quality *bit*; this was hypothesised to have been caused by the influence of spelling and the stringent role that spelling plays in Spanish.

Moreover, pronunciation training yielded significant results in priming students to start a phonemic categorisation process by producing different acoustic values for English high front/back tense and lax vowels.

In conclusion, this study shed light on the role of duration and spectra in the perception of non-native vowels in that it was shown that participants' awareness of basic perceptual contrasts can be *mis*used as phonemically contrastive cues.

4.4 Limitations and future research

This study encountered some limitations. The most important limitation was time constraints, and it affected the extent of this research and the number of participants. The second limitation that affected this research the most was the technical issue prevented the first group of participants from identifying certain stimuli.

With regard to indications for future research, given the insights provided by this research, research on the same age population of different L1 could yield potentially valuable insights related to the effect of native language on non-native cue weightings in vowel perception.

References

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, *114*(3), 1600–1610. https://doi.org/10.1121/1.1603234

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. In M. J. Munro & O. S. Bohn (Eds.), *Language experience in second language speech learning* (pp. 13–34). John Benjamins. https://doi.org/10.1075/lllt.17.07bes

Boersma, Paul & Weenink, David (2024). Praat: doing phonetics by computer [Computer program]. Version 6.4.12, retrieved 2 May 2024 from http://www.praat.org/

Bohn, O. S. (2019). Cross-language phonetic relationships account for most, but not all l2 speech learning problems: The role of universal phonetic biases and generalized sensitivities. In *Approaches to the Study of Sound Structure and Speech: Interdisciplinary Work in Honour of Katarzyna Dziubalska-Kolaczyk* (pp. 171–184). Taylor and Francis. https://doi.org/10.4324/9780429321757-13

Bohn, O. S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, *11*(3), 303–328. https://doi.org/10.1017/S0142716400008912

Bohn, O.-S. (1995). Cross Language Speech Perception in Adults First Language Transfer Doesn"t Tell It All. *Speech Perception and Linguistic Experience: Issues in Crosslanguage Research*.

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, *34*(3), 372–387. https://doi.org/10.1016/j.wocn.2005.08.003

Cerviño-Povedano, E., & Mora, J. C. (2015). Spanish EFL Learners' Categorization of /iː-I/ and Phonological Short-Term Memory. *Procedia - Social and Behavioral Sciences*, *173*, 18–23. https://doi.org/10.1016/j.sbspro.2015.02.024

Chung, H., & Kim, Y. (2021). Acoustic characteristics of Korean-English bilingual speakers' /l/ and the relationship to their foreign accent ratings. *Journal of Communication Disorders*, *94*. https://doi.org/10.1016/j.jcomdis.2021.106157

Corretge, Ramon (2024). Praat Vocal Toolkit. https://www.praatvocaltoolkit.com

(Crosshwhite, K. n.d). Praat Script.

ELvira-García, Wendy. 2021. A RShiny interface for plotting vowel formants. [RShiny App]

Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, *37*(4), 452–465. https://doi.org/10.1016/j.wocn.2009.07.006

Escudero, P., & Boersma, P. (2004). Bridging the gap between l2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*(4), 551–585. https://doi.org/10.1017/S0272263104040021

Farrelly, R. (2018). Lesson Planning for Teaching Pronunciation. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6). Wiley. https://doi.org/10.1002/9781118784235.eelt0242

Flege, J. E., Bohn, O. S., & Jang, S. (1997a). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437–470. https://doi.org/10.1006/jpho.1997.0052

Flege, J. E., Bohn, O. S., & Jang, S. (1997b). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437–470. https://doi.org/10.1006/jpho.1997.0052

Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3–83). Cambridge University Press. https://doi.org/DOI: 10.1017/9781108886901.002

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Article in The Journal of the Acoustical Society of America*, *76*(3), 702–721. https://www.researchgate.net/publication/353658738

Gottfried, T. L., & Beddor, P. S. (1988). Perception of temporal and spectral information in French vowels. *Language and Speech*, *31*(1), 57–75. https://doi.org/10.1177/002383098803100103

Hillenbrand, J., Getty, L. A., Wheeler, K., & Clark, M. J. (1994). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *95*(5_Supplement), 2875–2875. https://doi.org/10.1121/1.409456

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America*, *108*(6), 3013–3022. https://doi.org/10.1121/1.1323463

Morrison, G. S. (2003). Perception and Production of Spanish Vowels by English Speakers. In M. J. Solé, D. Recanses, & J. Romero (Eds.), *Proceedings of the 15th international congress of phonetic sciences* (pp. 1533–1536). Casual Productions.

Schertz, J., & Clare, E. J. (2020). Phonetic cue weighting in perception and production. In *Wiley Interdisciplinary Reviews: Cognitive Science* (Vol. 11, Issue 2). Wiley-Blackwell. https://doi.org/10.1002/wcs.1521

Souza, H. K. D., Carlet, A., Jułkowska, I. A., & Rato, A. (2017). Vowel inventory size matters: Assessing cue-weighting in L2 vowel perception. *Ilha Do Desterro*, *70*, 33–46. https://doi.org/https://doi.org/10.5007/2175-8026.2017v70n3p33

Wang, X. (2006). Mandarin listeners' perception of English vowels: Problems and strategies. *Canadian Acoustics*, *34*(4), 15–26. https://doi.org/http://hdl.handle.net/10211.3/179856

Wayland, R. (2018). *Phonetics: A practical introduction*. Cambridge University Press. https://doi.org/https://doi.org/10.1017/9781108289849

Yu, L. (2023). *The Perceptual Cue Weighting of English Tense-Lax Vowel Contrasts by First Language (L1) and Second Language (L2) Speakers of English* [Indiana University]. https://www.proquest.com/dissertations-theses/perceptual-cue-weighting-english-tense-lax-vowel/docview/2856634410/se-2?accountid=14733

Zhu, W., Lee, S. H., & Zhang, X. (2023). Language-dependent cue weighting in distinctive feature: evidence from the perception of Mandarin high vowels by native English speakers. *Asian-Pacific Journal of Second and Foreign Language Education*, *8*(1). https://doi.org/10.1186/s40862-023-00204-6

Appendix A. Consent Forms

Full de consentiment informat

Títol de l'estudi: Investigació de la percepció del contrast vocàlic de la segona llengua a l'aula — Un estudi sobre l'adquisició de les vocals laxes i tenses de l'anglès en parlants catalans/espanyols

Dades de contacte de l'investigador principal:

- Nom i cognoms: .... Adrián Salcedo Silva
- Correu electrònic: .... adriangerard.salcedo@alumni.urv.cat
- Telèfon: 977 55 83 82
- Adreça postal: .... Universitat Rovira i Virgili – Campus Catalunya – Facultat de Lletres – Av. Catalunya, 35 – 43002 Tarragona

Jo .... amb DNI ....

- He llegit el full d'informació al participant sobre l'estudi del qual se m'ha entregat una còpia.
- He pogut fer preguntes i resoldre els meus dubtes sobre l'estudi i la meva participació.
- Comprenc la meva participació a l'estudi d'acord amb allò expressat al full d'informació al participant sobre l'estudi i de les respostes a les meves preguntes, així com els riscos i beneficis que comporta.
- Accepto que la meva participació és voluntària i dono lliurement la meva conformitat per participar a l'estudi.
- Conec que em puc retirar en qualsevol moment de la participació a l'estudi sense que això em pugui causar cap perjudici.
- Estic informat sobre el tractament que es realitzarà de les meves dades personals.
- Dono el meu consentiment per a l'accés i utilització de les meves dades en les condicions detallades al full d'informació al participant sobre l'estudi.

  ☐**Sí** ☐**No**
- Dono el meu consentiment per a la difusió de les meves dades personals junt amb la publicació dels resultats de l'estudi.

  ☐**Sí** ☐**No**
- Un cop finalitzada la investigació, és possible que les dades obtingudes siguin d'interès per a altres estudis relacionats. En relació amb això, s'ofereixen les següents opcions:

  ☐**NO autoritzar** l'ús de les seves dades en altres projectes d'investigació relacionats.

  ☐**SÍ autoritzar** l'ús de les seves dades en altres projectes d'investigació relacionats.
- Un cop finalitzada la investigació, és possible que hi hagi mostra sobrant. En relació a aquestes, s'ofereixen les següents opcions:

  ☐La **destrucció** de la mostra sobrant.

  ☐La seva **utilització en futurs projectes** d'investigació biomèdica relacionats amb el mateix tema

I per expressar aquest consentiment, el participant signa en data .... i lloc .... aquest full de consentiment:

Signatura del participant..

MODEL DE FULL D'INFORMACIÓ AL PARTICIPANT

TÍTOL DE L'ESTUDI

Investigació de la percepció del contrast vocàlic de la segona llengua a l'aula — Un estudi sobre l'adquisició de les vocals laxes i tenses de l'anglès en parlants catalans/espanyols

INVESTIGADOR PRINCIPAL / DOCTORAND / ESTUDIANT

Adrián Salcedo Silva – adriangerard.salcedo@alumni.urv.cat – Telèfon: 977 55 83 82 – Adreça postal: Universitat Rovira i Virgili – Campus Catalunya - Facultat de Lletres – Av. Catalunya, 35 – 43002 Tarragona

CENTRE

Universitat Rovira i Virgili (URV) – Facultat de Lletres

INTRODUCCIÓ

Ens dirigim a vostè per tal d'informar-lo sobre l'estudi d'investigació en el que se'l convida a participar. Aquest estudi ha estat aprovat pel Comitè Ètic d'investigació en persones, societat i medi ambient **de la Universitat Rovira i Virgili.**

La nostra intenció és que rebi la informació correcta i suficient perquè pugui avaluar i decidir si vol o no participar en aquest estudi. Per aquest motiu, llegeixi aquest full informatiu amb atenció i nosaltres li aclarirem els dubtes que li puguin sorgir Addicionalment, li informem que vostè es lliure de consultar amb les persones que consideri oportú abans de decidir sobre la seva participació a l'estudi.

PARTICIPACIÓ VOLUNTÀRIA

Ha de saber que la seva participació en aquest estudi és voluntària i que pot decidir no participar o canviar la seva decisió i retirar el consentiment en qualsevol moment.

DESCRIPCIÓ GENERAL DE L'ESTUDI

Aquest estudi té com a objectiu comprovar la millora en la percepció i la pronunciació de vocals angleses de parlants catalans/espanyols.

L'estudi portarà a terme tres intervencions metodològiques per ensenyar la pronunciació de l'anglès a estudiants d'educació secundària amb l'objectiu de fomentar la correcta percepció i articulació dels sons anglesos. Les intervencions es basaran en la duració de les vocals, la qualitat de les vocals, i una metodologia implícita de repetició.

No serà necessària cap visita mèdica; només caldrà que respongui el qüestionari amb preguntes relacionades amb l'estil de vida previ a l'estudi.

BENEFICIS I RISCOS

Aquest estudi ajudarà a esclarir la relació entre qualitat i duració en la percepció i producció de vocals en l'adquisició dels sons anglesos. També ajudarà a aportar informació sobre l'eficàcia de metodologies d'ensenyament de la pronunciació.

L'estudi no suposa cap risc per al participant.

CONFIDENCIALITAT I PROTECCIÓ DE DADES

Tota la informació recopilada sobre les persones participants en el marc d'aquest estudi es mantindrà estrictament confidencial i amb aplicació de les corresponents mesures de seguretat que garanteixin, a més de la seva confidencialitat, la seva integritat, disponibilitat, autenticitat i traçabilitat.

Les dades personals recollides per a l'estudi estaran identificades mitjançant un codi i només l'investigador principal o els seus col·laboradors podran relacionar aquestes dades amb els participants. Mai s'identificarà a les persones participants en cap informe, presentació ni publicació que sorgeixi d'aquest estudi. Per tant, la seva identitat no serà revelada a cap persona, excepte quan sigui requerit pel Comitè d'Ètica al que es sotmet l'estudi amb la finalitat de comprovar les dades i procediments de l'estudi.

Per al tractament de les dades s'utilitzaran els sistemes d'informació propis de la Universitat Rovira i Virgili instal·lats a la seva xarxa informàtica aplicant-se les mesures de seguretat de la informació establertes pel Reial Decret 3/2010 que regula l'Esquema Nacional de Seguretat. Concretament, les dades es recolliran mitjançant gravacions de veu i s'introduiran en el sistema d'informació Microsoft OneDrive de la URV. Posteriorment, per analitzar les dades s'utilitzarà el programa Praat i Jasp.

El personal investigador de l'estudi es compromet a complir la Llei orgànica 3/2018, de 5 de desembre, de protecció de dades personals i garantia dels drets digitals, a més del Reglament (UE) núm. 2016/679, del Parlament europeu i del Consell, de 27 d'abril de 2016, relatiu a la protecció de les persones físiques pel que fa al tractament de dades personals, i signarà un compromís de participació i confidencialitat.

La finalitat del tractament de les dades és la participació en l'estudi d'acord amb el consentiment del seu tutor legal. La persona participant també pot donar el consentiment per a la reutilització de les dades per a estudis futurs que estiguin relacionats.

La persona participant podrà interrompre la seva participació a l'estudi o estudis futurs relacionats retirant el seu consentiment en qualsevol moment, sense que sigui necessària la seva justificació. En aquest cas, les dades no es podran eliminar per tal de garantir la validesa dels resultats i complir amb les obligacions legals aplicables a l'estudi, però sí que quedaran codificades de manera que no sigui possible vincular-les a la seva persona.

INFORMACIÓ AMPLIADA SOBRE EL TRACTAMENT DE DADES PERSONALS

De conformitat amb el que disposa la legislació vigent en matèria de protecció de dades aplicable a la Universitat Rovira i Virgili (URV) i publicada a l'apartat "Legislació aplicable" de l'espai "Protecció de dades de caràcter personal" de la Seu Electrònica (https://seuelectronica.urv.cat/rgpd/), es posa en coneixement de les persones interessades la informació següent:

a) Qui és el responsable del tractament de les seves dades?

• **Identificació**	Universitat Rovira i Virgili CIF: Q9350003A
• **Adreça Postal**	Carrer de l'Escorxador, s/n 43003 Tarragona
• **Dades de contacte dels DPD**	DPD - Delegats de protecció de dades de la URV Correu electrònic: dpd@urv.cat

b) Quines dades personals tractem i amb quina finalitat?

Les dades personals són tractades amb la finalitat de participar en l'estudi del treball de fi de màster en els termes que es descriuen al full d'informació al participant. En el cas que l'estudi prevegi la publicació, difusió i reutilització dels resultats obtinguts incloent dades personals, les dades personals seran utilitzades per a aquesta finalitat sempre que l'interessat hagi atorgat el seu consentiment.

c) A quins destinataris es comunicaran les seves dades?

En el marc del tractament mencionat, les seves dades no es cediran a tercers tret que existeixi obligació legal o s'indiqui expressament en el full d'informació al participant.

d) Quina és la legitimació per al tractament de les seves dades?

La legitimació d'aquest tractament es basa en el consentiment que dona la persona interessada de forma expressa.

e) Quines mesures de seguretat apliquem en el tractament de les seves dades?

La Universitat es responsabilitza d'aplicar les mesures de seguretat i la resta d'obligacions derivades de la legislació de protecció de dades de caràcter personal, d'acord amb l'Esquema Nacional de Seguretat, Reial Decret 3/2010.

En aquest sentit, La Universitat Rovira i Virgili s'ha dotat d'una Política de Seguretat que pot ser consultada a la secció sobre "Legislació i normativa" de la pàgina web de la Universitat dintre de "Normativa pròpia" i "Altres normes", http://www.urv.cat/ca/universitat/normatives/altres-normes/.

Addicionalment, al Full d'informació al participant es concreten algunes mesures de seguretat específiques que es tindran en compte durant la realització de l'estudi.

f) Quins són els drets dels interessats?

L'interessat té dret a accedir a les seves dades personals; a demanar la rectificació de les dades inexactes; a sol·licitar la cancel·lació i supressió; a oposar-se al tractament, inclosa l'elaboració de perfils;

a limitar fins a una data determinada el tractament de les seves dades; i a la portabilitat de les mateixes en format electrònic.

La persona participant pot interrompre la seva participació a l'estudi retirant el seu consentiment en qualsevol moment, sense donar explicacions. En aquest cas, les dades no es podran eliminar per tal de garantir la validesa dels resultats i complir amb les obligacions legals aplicables a l'estudi, però no serà possible vincular-les a la seva persona.

Podrà exercir els drets d'accés, rectificació, cancel·lació, oposició, limitació i portabilitat mitjançant comunicació escrita, detallant motivadament la sol·licitud, adreçada al Registre General (Carrer de l'Escorxador, s/n, 43003 de Tarragona) o mitjançant la seva presentació al Registre General de la Universitat, presencialment o telemàtica, segons s'indica a https://seuelectronica.urv.cat/registre.html.

Així mateix, l'informem que té dret a presentar una reclamació davant l'Autoritat Catalana de Protecció de Dades mitjançant el mecanisme que estableixi. Pot consultar més informació a https://apdcat.gencat.cat/ca/inici.

Finalment, l'informem que podrà sol·licitar informació relacionada amb la protecció de dades personals mitjançant correu electrònic als nostres delegats de protecció de dades a la direcció del dpd@urv.cat.

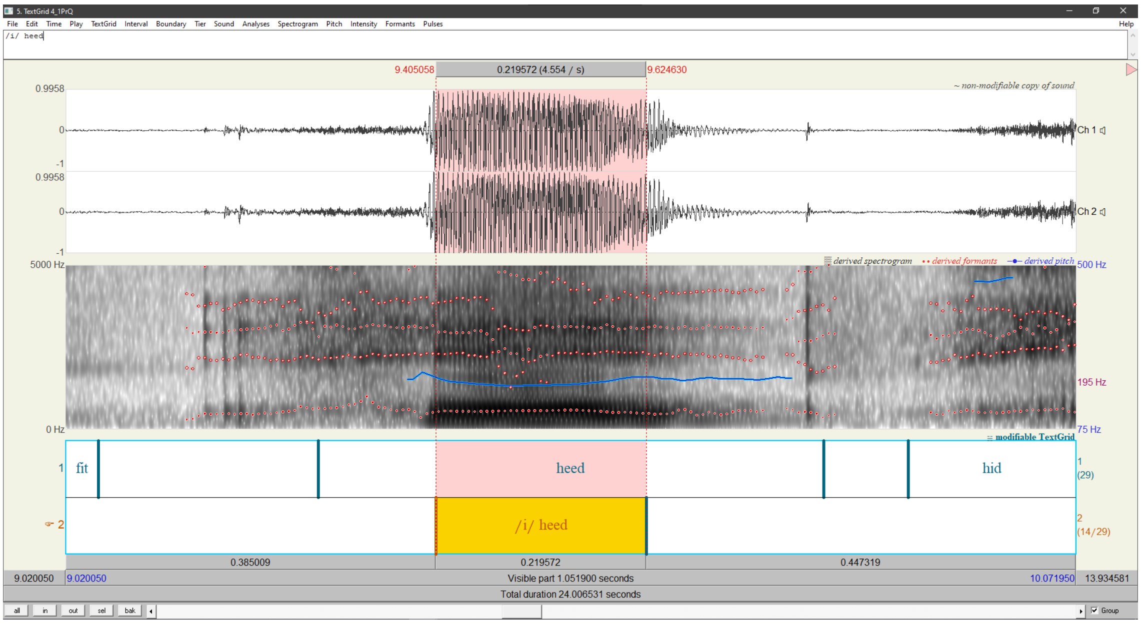
g) Quant de temps conservarem les seves dades?

El període de conservació de les dades és de 5 anys un cop finalitzat l'estudi, tret que el full d'informació al participant estableixi un període diferent. En qualsevol cas, es conservaran les dades fins a la revocació del consentiment per part de la persona interessada.

MOSTRES A RECOLLIR

Gravació de veu.

Appendix B. Praat materials



File Edit Search Convert Font Run

```
###Description of this script
##  This script measures f0 and the first three formants at the midpoint of the vowel, and appends the
##  results to a text file.  It will be called "formant-log.txt", and will be written to the same
##  directory holding your sound files.
##  To run this script, you will have to have a bunch of sound files with accompanying text grids.  The
##  locations of vowels to be measured must be marked in tier 1 of the textgrid.  Anything with a non-null
##  label in that tier will be logged.
###End of description

##  Specify the directory containing your sound files in the next line:
```

Appendix C. Response per stimulus

Table 1. Total "beat" responses per stimulus (% *beat* id.) for IT in pre-test and post-test

	L1		L2		L3		L4		L5	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Q1	15	17	15	19	12	23	9	18	9	16
	(50)	(57)	(50)	(63)	(40)	(77)	(30)	(60)	(30)	(53)
Q2	19	20	—	23	14	15	14	22	8	14
	(63.3)	(67)	—	(77)	(46.7)	(50)	(46.7)	(73)	(26.7)	(47)
Q3	24	23	21	22	18	13	—	19	9	17
	(80)	(77)	(70)	(73)	(60)	(43)	—	(63)	(30)	(57)
Q4	22	23	20	19	—	17	17	20	—	16
	(73.3)	(77)	(66.7)	(63)	—	(57)	(56.7)	(67)	—	(53)
Q5	21	18	23	18	15	17	18	12	13	17
	(70)	(60)	(76.7)	(60)	(50)	(57)	(60)	(40)	(43.3)	(57)

Table 2. Total "beat" responses per stimulus (% *beat* id.) for QT in pre/post-test

	L1		L2		L3		L4		L5	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Q1	9	12	11	10	9	8	5	9	5	8
	(45)	(60)	(55)	(50)	(45)	(40)	(25)	(45)	(25)	(40)
Q2	13	10	11	11	13	7	9	10	8	10
	(65)	(50)	(55)	(55)	(65)	(35)	(45)	(50)	(40)	(50)
Q3	10	6	15	10	13	10	12	11	10	9
	(50)	(30)	(75)	(50)	(65)	(50)	(60)	(55)	(50)	(45)
Q4	14	10	14	7	8	5	9	11	12	10
	(70)	(50)	(70)	(35)	(40)	(25)	(45)	(55)	(60)	(50)
Q5	15	15	12	12	13	11	14	10	10	14
	(75)	(75)	(60)	(60)	(65)	(55)	(70)	(50)	(50)	(70)

Table 3. Total "beat" responses per stimulus (% *beat* id.) for LT in pre-test and post-test

	L1		L2		L3		L4		L5	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Q1	20 (55.6)	21 (58.3)	19 (52.8)	21 (58.3)	16 (44.4)	27 (75)	18 (50)	20 (55.6)	22 (61.1)	13 (36.1)
Q2	18 (50)	22 (61.1)	22 (61.1)	25 (69.4)	22 (61.1)	25 (69.4)	17 (47.2)	18 (50)	22 (61.1)	23 (63.9)
Q3	23 (63.9)	18 (50)	22 (61.1)	20 (55.6)	18 (50)	18 (50)	24 (66.7)	23 (63.9)	20 (55.6)	17 (47.2)
Q4	20 (55.6)	21 (58.3)	24 (66.7)	19 (52.8)	16 (44.4)	14 (38.9)	17 (47.2)	20 (55.6)	20 (55.6)	19 (52.8)
Q5	22 (61.1)	23 (63.9)	22 (61.1)	18 (50)	18 (50)	24 (66.7)	20 (55.6)	18 (50)	22 (61.1)	15 (41.7)

Table 4. Total number of "beat" responses per stimulus (% *beat* id.) for IT group in pre-test

	L1	L2	L3	L4	L5
Q1	4 (100)	4 (100)	4 (100)	4 (100)	3 (75)
Q2	3 (75)	3 (75)	2 (50)	2 (50)	2 (50)
Q3	2 (50)	0 (0)	1 (25)	0 (0)	0 (0)
Q4	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Q5	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)