UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO INGLÊS: ESTUDOS LINGUÍSTICOS E LITERÁRIOS

Elisabeth Ann Bunch Oliveira da Rosa

The effect of perception training with synthetic and natural stimuli on BP learners' ability to identify the English vowels /æ-ɛ/

Florianópolis

2022

Elisabeth Ann Bunch Oliveira da Rosa

The effect of perception training with synthetic and natural stimuli on BP learners' ability to identify the English vowels /æ-ɛ/

Dissertação submetida ao Programa de Pós-graduação em Inglês da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Inglês: Estudos Linguísticos e Literários.
Orientadora: Profa. Dra. Rosane Silveira

Florianópolis

2022

Ficha de identificação da obra

Rosa, Elisabeth Ann Bunch Oliveira da
The effect of perception training with synthetic and natural stimuli on BP learners' ability to identify the English vowels /æ-?/ / Elisabeth Ann Bunch Oliveira da Rosa ; orientador, Rosane Silveira, 2022.
95 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Programa de Pós Graduação em Inglês: Estudos Linguísticos e Literários, Florianópolis, 2022.

Inclui referências.

1. Inglês: Estudos Linguísticos e Literários. 2. Percepção. 3. Vogais inglesas. 4. Estímulos sintéticos. 5. Peso de pistas acústicas. I. Silveira, Rosane. II. Universidade Federal de Santa Catarina. Programa de Pós Graduação em Inglês: Estudos Linguísticos e Literários. III. Título.

Elisabeth Ann Bunch Oliveira da Rosa

The effect of perception training with synthetic and natural stimuli on BP learners' ability to identify the English vowels /æ-ɛ/

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof.ª Hanna Emilia Kivistö-de Souza, Dr.ª
UFSC

Prof.ª Angélica Carlet, Dr.ª
UIC

Prof. Alison Roberto Gonçalves, Dr.
UFPR

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Inglês: Estudos Linguísticos e Literários.

Documento assinado digitalmente
Celso Henrique Soufen Tumolo
Data: 06/09/2022 14:00:51-0300
CPF: 030.275.298-62
Verifique as assinaturas em https://v.ufsc.br

Coordenação do Programa de Pós-Graduação

Documento assinado digitalmente
Rosane Silveira
Data: 06/09/2022 12:00:51-0300
CPF: 715.936.839-87
Verifique as assinaturas em https://v.ufsc.br

Prof.(a) Rosane Silveira, Dr.(a)
Orientador(a)

Florianópolis, 2022.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the help of many others along the way. I would like to thank the following people.

First, thank you to my advisor, Professor Rosane Silveira. Thank you so much for your patience and guidance throughout this project. I have learned so much from you. I admire the wisdom and care you have for pronunciation teaching. The knowledge you shared in the Learning, Teaching, and Assessing Pronunciation class that I took from you my first semester at UFSC greatly impacted some of my beliefs about the ethics of English teaching, and I will be forever grateful.

Thank you to Valdete and all of the secretarial staff in PPGI for your continued support throughout the program, for your patience and kindness, and for the welcoming and encouraging environment that you provide in the office. Thank you to Professor Celso Tumolo for helping me through the first steps of enrollment. Thank you to all of the professors and staff that were a part of my journey through UFSC. I would also like to thank CAPES for the financial support throughout my master's program.

Thank you, Robson, for taking me under your wing into the phonetics and phonology world. Thank you for your friendship, for your contagious passion and excitement about the field, and for sharing your knowledge and experiences with me. Thank you to Brenda, Carol, Luana, and Vanessa, for being some of my first friends here in Brazil! Thank you, along with all of our other classmates, for welcoming me so warmly and for helping me throughout the Master's program.

I would like to thank my *Inglês Na Hora* students and all of those who are a part of our community. Your love for learning inspires me.

I would like to thank my parents, Jerry and Kim. Thank you, Dad, for passing on to me your curiosity and love for learning. Seeing you go back to school to get your degree as an adult, while juggling many other responsibilities simultaneously impressed upon me the value of education and encouraged me to continue furthering my own education as well. Thank you, Mom, for teaching me to be resilient, to believe in myself, and to prioritize what is important. Often when I feel overwhelmed, I remember your wise words to me from History Day in Mr. Hart's class.

Thank you to my loving husband, Léo. I wouldn't have been able to finish my degree without you. Thank you for helping me with all of my documentation arriving here in Brazil and getting enrolled at UFSC. Thank you for all the times you helped me to translate and explained unfamiliar vocabulary in Portuguese. Thank you for staying up late with me while I studied, for spending a full week collecting data with me, and for taking on more responsibilities in work and around the house so that I could have more time to study. Thank you for encouraging me and loving me well when I got stressed. Te amo.

ABSTRACT

Perception is a crucial component in the acquisition of a second language (L2) regarding oral communication. Research has revealed factors that can often predict the specific difficulties for acquiring certain sounds in the L2. Vowels tend to pose a special difficulty, with the English vowel pair /æ-ɛ/ being particularly difficult for native Brazilian Portuguese (BP) learners, who may not distinguish the two as separate, but instead perceive them both as the vowel /ɛ/. Perception training with synthetic stimuli is one way to assist L2 learners in the formation of new vowel categories. Synthetic stimuli that control for vowel duration may be especially effective for this type of training, as it assists learners in developing appropriate cue weighting ability for the L2. Based on the discussion above, the present study investigates the effectiveness of synthetic versus natural stimuli for perception training on the ability of Brazilian learners of English to identify the vowels /æ-ɛ/. The participants for this study were 56 Brazilian learners of English, divided into natural stimuli, synthetic stimuli, and control groups. Participants received perception training through the online platform Gorilla and completed pre and post training perception tests to measure their progress. Results indicate that perception training with both synthetic and natural stimuli is effective for the identification of /æ-ɛ/.

Keywords: 1. Perception 2. English vowels 3. Synthetic stimuli 4. Cue weighting

RESUMO

A percepção é um componente crucial na aquisição de uma segunda língua no que diz respeito à comunicação oral. Pesquisas revelaram fatores que costumam prever as dificuldades específicas para adquirir certos sons na L2. As vogais tendem a apresentar uma dificuldade especial, com o par de vogais em inglês /æ-ɛ/ sendo particularmente difícil para os alunos nativos do português brasileiro, que podem não distinguir os dois como categorias separadas, mas sim percebê-los como a vogal /ɛ/. O treinamento de percepção com estímulos sintéticos é uma maneira de ajudar os aprendizes de segunda língua na formação de novas categorias de vogais. Estímulos sintéticos que controlam a duração da vogal e manipulam os valores de F1 e F2 podem ser especialmente eficazes para esse tipo de treinamento, pois ajudam os aprendizes a desenvolver a capacidade de utilizar as pistas acústicas de uma maneira adequada para a segunda língua. Com a base no acima exposto, a presente pesquisa investiga a eficácia dos estímulos sintéticos versus naturais para o treinamento de percepção e para o desenvolvimento da capacidade de aprendizes brasileiros do inglês em distinguir o par de vogais /æ-ɛ/**.** Os participantes deste estudo são aprendizes brasileiros de inglês, 24 iniciantes e 24 avançados, divididos em grupos de estímulos naturais, estímulos sintéticos e controle. Os participantes receberam treinamento de percepção com o auxílio da plataforma online Gorilla e concluíram testes de percepção pré e pós-treinamento para medir seu progresso. Os resultados indicam que o treinamento perceptual com estímulos sintéticos e naturais auxilia na identificação das vogais /æ-ɛ/.

Keywords: 1. Percepção 2. Vogais inglesas 3. Estímulos sintéticos 4. Peso de pistas acústicas

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

L2 Second language

L1 Native language

BP Brazilian Portuguese

SLM Speech Learning Model

SLM-r Revised Speech Learning Model

ID Identification

DIS Discrimination

F1 First formant

F2 Second formant

F3 Third formant

F4 Fourth formant

F5 Fifth formant

ELF English as a Lingua Franca

RQ Research question

H Hypothesis

TABLE OF CONTENTS

1 INTRODUCTION

Perception is a crucial component in the acquisition of a second language (L2) regarding oral communication. Adult learners' perceptions of speech is typically patterned by the phonemic inventory of his or her native language (L1) (FLEGE, 1995). Research has revealed factors that can often predict the difficulty for acquiring specific sounds in the second language. The Speech Learning Model, for example, does this based on phonetic similarity (FLEGE, 1995). This model suggests that an L2 sound that is perceived as being similar to an L1 category will be more difficult to acquire (FLEGE, 1995) than a sound that is perceived as being more different. In the case of native Brazilian Portuguese (BP) speakers learning English, this would include vowel pairs such as /æ-ɛ/, as the vowel /ɛ/ exists in the BP vowel inventory but the vowel /æ/, while perceived as being similar to /ɛ/ (NOBRE-OLIVEIRA, 2007), does not. BP English learners may therefore perceive these pairs as the same vowel (for example, perceiving both /æ-ɛ/ as /ɛ/). Research has shown that the vowel pair /æ-ɛ/ is in fact one of the most challenging for BP learners of English to perceive in comparison to other vowel pairs, typically assimilating both /æ-ɛ/ vowels into the BP category /ɛ/ (LIMA JR., 2017; RAUBER, 2006).

Perception training with synthetic stimuli is one way to assist L2 learners in the formation of new vowel categories. Synthetic stimuli can be created by manipulating a speech sound. Ladefoged (2007) identifies nine components for a speech sound, namely:

> (1-3) the frequencies of the first three formants (F1, F2, F3), (4-6) the amplitudes of the first three formants, (7-8) the frequency and amplitude of the voiceless components, and (9) the fundamental frequency of voiced sounds (LADEFOGED, 2007, p. 68).

Synthetic stimuli refer to the manipulation of one or more of these components. In this study, the duration of the vowel sounds within carrier words will be manipulated to create synthetic stimuli.[1]

Research has shown that synthetic stimuli can be just as, or possibly even more effective than natural stimuli, as with synthetic stimuli, the "subtle and crucial cues of the [acoustic] signal are enhanced, drawing learners' attention to them (and the less important features

---

[1] Section 2.3 of this thesis will discuss relevant acoustic terms in detail, especially those that are important for the formation of vowel categories, such as F2.

attenuated)" (NOBRE-OLIVEIRA, 2007, p. 148). Cheng, Zhang, Fan, and Zhang's study on temporal exaggeration also supports the use of acoustic exaggeration in the high variability phonetic training method (2019).

Although there have been various studies on the effectiveness of perception training (ALVES; LUCHINI, 2017; CARLET; CEBRIAN, 2014; CHENG; ZHANG; FAN; ZHANG, 2019; NOBRE-OLIVEIRA, 2007; RATO, 2014; STRANGE; DITTMAN, 1984; YLINEN; UTHER; LATVALA; VEPSALAINEN; IVERSON; AKAHANE-YAMADA; NAATANEN, 2009) the effects of natural vs. synthetic stimuli (CHENG ET. AL, 2019; NOBRE-OLIVEIRA, 2007) and the ability of learners to generalize their learning to new stimuli and speakers (CARLET; CEBRIAN, 2014; NOBRE-OLIVEIRA, 2007), there have been few studies comparing the effects of perception training with natural versus synthetic stimuli for learners of different proficiency levels. The present study aims to examine the use of natural versus synthetic stimuli on the perception of the vowel pair /æ-ɛ/ and also the effectiveness of this training for learners at different proficiency levels.

1.1 SIGNIFICANCE OF THE STUDY

L2 pronunciation and perception is an important part of second language acquisition; however, often teachers feel inadequately equipped to provide students with proper instruction in the classroom. This is sometimes due to a lack of suitable teaching and learning materials that are of good quality for the L2 classroom (COSTA, 2016; MACDONALD, 2002). As such, it is crucial to provide L2 teachers with effective resources, enabling them with the confidence and skill set necessary to provide instruction for their L2 students. The results of this study may be beneficial to L2 teachers and L2 programs, as it will indicate the effectiveness of perception training with natural and synthetic stimuli for L2 learners. If this training proves to be successful for participants' ability to distinguish the English vowel pair /æ-ɛ/, it may encourage teachers to utilize this method in their classrooms. If this study indicates the unsuccessfulness of perception training for this vowel pair, it may promote further research in other methods for L2 perception development. Perceptual training may also ease oral production (QIAN; CHUKHAREV-HUDILAINEN; LEVIS, 2018), and therefore this study is significant for L2 learners' perception and speech. In addition to the relevance for L2 instructional programs, this study will contribute to the growing knowledge base of L2 pronunciation and perception development in general.

1.2 ORGANIZATION OF THE PROJECT

Having outlined the context of investigation and briefly presented the research problem, the following section describes the main objective and specific objectives for the present study (Section 1.2). A review of the literature will follow (Chapter 2) to provide a framework for the present study and to summarize other studies that have been done in this area. The method to be used for this study will then be explained in Chapter 3, detailing the participants, instruments for data collection, procedures for data collection, and procedures for data analysis. Chapter 4 will follow with the results and discussion. The thesis will conclude with some final remarks, a discussion of the limitations of the study, and recommendations for further research in Chapter 5.

1.3 OBJECTIVES

This section will describe both the main objective and specific objectives for this study regarding the effectiveness of natural and synthetic stimuli.

1.3.1 Main objective

The main objective of this study is to understand the effectiveness of natural and synthetic stimuli for perception training with the English vowels /æ-ɛ/ for native Brazilian-Portuguese L2 English students at various L2 proficiency levels.

1.3.2 Specific objectives

The specific objectives for this study will be

- To compare the effect of natural vs. synthetic stimuli for perception training with the vowel pair /æ-ɛ/.
- To understand the ability of students to generalize the gains from their perception training with natural vs. synthetic stimuli to natural stimuli with new voices.
- To understand whether L2 learners are able to maintain their learning from perception training with natural vs synthetic stimuli
- To compare the effects of natural stimuli and synthetic stimuli for perception training based on proficiency level.

In order to explore these objectives, this study will answer four research questions (RQ):

RQ1: What is the effect of natural versus synthetic stimuli for perception training with the vowel pair /æ-ɛ/?

RQ2: How successfully are BP L2 English students able to generalize their perception training with natural versus synthetic stimuli to unfamiliar voices and carrier words?

RQ3: How successfully are BP L2 English students able to maintain the effects of perception training with natural versus synthetic stimuli?

RQ4: How does proficiency level impact the effectiveness of identification perception training for BP L2 English students?

Based on previous studies on the use of synthetic and natural stimuli for perception training (CHENG ET. AL, 2019; NOBRE-OLIVEIRA, 2007; YLINEN ET.AL, 2009; WANG ET. AL, 1999) the following hypotheses (H) were developed in response to each of the research questions listed above.

H1: BP L2 English students that complete perception training with synthetic stimuli will show more improvement in their ability to identify the English vowels /æ-ɛ/ than those that complete perception training with natural stimuli. Previous research has shown that both synthetic and natural stimuli are effective in perception training. (CHENG ET. AL, 2019; NOBRE-OLIVEIRA, 2007). However, in Cheng et. al's 2019 study, it was shown that participants who trained with synthetic stimuli were more successful in generalizing their learning to new speakers and words. In Nobre-Oliveira's 2007 study, the data indicated that synthetic stimuli may be more effective for perception training, although the data was not statistically significant.

H2: BP L2 English students that complete perception training with synthetic stimuli will be better at generalizing their learning to new voices and carrier words than those who complete perception training with natural stimuli. As mentioned above, Cheng et. al's 2019 study specifically found a difference in participants' ability to generalize their learning based on the type of stimuli they used for perception training. In addition to this, Escudero et. al notes that native speakers tend to rely more on formant frequency cues than they do on durational cues when identifying English vowels (2004). Considering then that all of the synthetic stimuli in this study will have a controlled duration of 300ms (thus preventing the ability for participants to depend on durational cues,) it is hypothesized that these participants will learn to depend more consistently on formant frequency differences and in turn identify the English vowels in new voices and new carrier words with a higher accuracy.

H3: BP L2 English students that complete perception training with synthetic stimuli will be able to maintain their learning at a higher level than those that complete perception training with natural stimuli. This hypothesis is based on results from Nobre-Oliveira's 2007 study, which showed that (although the results were not statistically significant,) while participants who trained with natural stimuli maintained their learning, with the same score from post-test to follow-up test, participants who trained with synthetic stimuli improved from post-test to follow-up test.

H4: BP L2 English students with a high L2 proficiency level will benefit more from perception training than those with a low L2 proficiency level. This hypothesis is based on the role that orthography plays in L2 speech perception (STOHER; MARTIN, 2021.) Students with a high L2 proficiency level may be less likely to deal with spelling or low-frequency word challenges than students with a low L2 proficiency.

2 REVIEW OF THE LITERATURE

This chapter will first review the Speech Learning Model proposed by Flege (1995), which will serve as the theoretical background for the present study. Following a review of this model will be an explanation of phonetic training in general, and then a discussion about two primary methods for perception training. Studies which have explored the effectiveness of perception training and the effectiveness of natural versus synthetic stimuli for perception training will then be discussed. Finally, the differences between the BP and General American English vowel inventories will be explored, noting the specific difficulties that BP learners of English typically encounter regarding vowel perception.

2.1 SPEECH LEARNING MODEL

As mentioned in the previous section, this research is focused on the effect of natural versus synthetic stimuli on beginner and advanced Brazilian Portuguese (BP) speakers of English's ability to identify the English vowels /æ-ɛ/. From the current L2 speech learning models – the Perceptual Assimilation Model-L2 (BEST; TYLER, 2007), the Native Language Magnet Theory (KUHL; CONBOY; COFFEY-CORINA; PADDEN; RIVERA-GAXIOLA; NELSON, 2007), and the Speech Learning Model (FLEGE, 1995) - the most appropriate for this project is the Speech Learning Model (SLM), as this model predicts the learning of new L2 categories, and as this project will deal with identification training and perception tests. Thus, only the SLM will be reviewed in this project.

SLM was originally developed by Flege in 1995, however in 2021, an updated version, the Revised Speech Learning Model (SLM-R) was developed by Flege and Bohn. The primary aim of the SLM is to "account for how phonetic systems reorganize over the life-span in response to the phonetic input during naturalistic L2 learning" (FLEGE; BOHN, 2021, p.23). The main assumptions of this model are that 1) "the phonetic categories which are used in word recognition and to define the targets of speech production are based on statistical input distributions," 2) "L2 learners of any age make use of the same mechanisms and processes to learn L2 speech that children exploit when learning their L1," and 3) "native versus nonnative differences in L2 production and perception are ubiquitous not because humans lose the capacity to learn speech at a certain stage of typical neuro-cognitive development but because applying the mechanisms and processes that functioned 'perfectly' in L1 acquisition to the sounds of an L2 do not yield the same results" (FLEGE; BOHN, 2021, p.23). The SLM provides an explanation for the differences in L1 and L2 learning outcomes, noting that in the early stages of learning, L1 sounds "substitute" L2 sounds, and these existing L1 phonetic categories can block new category formation for L2 sounds (FLEGE et. al., 2021). Flege explains that new category formation can be predicted based on the perceived phonetic dissimilarity of the L2 speech sounds from the closest L1 sounds, category formation being more likely for those that are more dissimilar (FLEGE, 2003). Flege also notes that "if instances of an L2 speech sound category persist in being identified as instances of an L1 speech sound, category formation for the L2 speech sound will be blocked" (2003, p. 10).

It can be predicted then, following the SLM propositions, that Brazilian learners will have more difficulty forming new categories for English phonemes that are perceived as being similar to phonemes in the Brazilian Portuguese inventory versus forming categories for English phonemes that are perceived as very different from those that are existent in their L1. It can also be predicted that to cope with this difficulty, Brazilian learners will merge the L1 and L2 categories that are perceived as being similar and have difficulty distinguishing them as separate phonemes.

Research supports the merging of L1 and L2 categories for Brazilian learners of English with the vowel pair /æ-ɛ/ (NOBRE-OLIVEIRA, 2007; RAUBER, 2006). The phoneme /ɛ/ exists in the Brazilian Portuguese phonetic inventory, however /æ/ does not. In accordance with SLM, Nobre-Oliveira's study found that /æ/ was misidentified as /ɛ/ most of the time (NOBRE-OLIVEIRA, 2007). Another study found that the vowel pair /æ-ɛ/ was the most poorly perceived pair (compared to /i-ɪ/ and /u-ʊ/) for BP learners of English (RAUBER, 2006). These results suggest that BP learners perceive /æ/ as being similar to /ɛ/ and treat both phonemes as a single category in English, as predicted by the SLM. The SLM therefore would predict difficulty for BP learners of English in acquiring the ability to perceive the /æ/ phoneme, as it is similar to the /ɛ/ phoneme.

2.2 PHONETIC TRAINING

Research has shown that perception training that utilizes repeated exposure to the key sounds is effective in improving learners' L2 perception. This has been successful with the distinction between the English vowel pair /i-ɪ/ for Chinese learners of English (CHENG ET. AL, 2019; WANG; MUNRO, 1999), the identification of English /i ʌ b v d/ by bilingual

Catalan-Spanish speakers (CARLET ET. AL, 2014), the identification of English /i-ɪ/ by Finnish users of English (YLINEN ET. AL 2009), the perception (and production) of word-initial voiceless stops by Argentinean learners of English (ALVES ET. AL, 2017), the discrimination between the English liquid consonants /r/ and /l/ by Japanese learners of English (STRANGE ET. AL, 1984), and the English vowel contrasts /i-ɪ/, /æ-ɛ/, or /u-ʊ/ by European Portuguese learners (RATO, 2014).

In high variability phonetic training (HVPT), listeners are specifically exposed to multiple voices speaking target sounds or words, and the listeners must identify or discriminate between the sounds that are presented. This method was originally developed by Logan, Lively, and Pisoni in a 1991 study. Logan et. al (1991) trained native speakers of Japanese learning English to identify the L2 sounds /r/ and /l/. The researchers emphasized variability in the stimuli, unique from previous studies up until that time. Results indicated that participants who received this training with high variability in the stimuli improved in their ability to identify the L2 sounds /r/ and /l/, whereas the participants who did not receive this training did not. Logan et. al found that this variability in voices helped listeners to disregard speech differences that were irrelevant to the perception of an L2 contrast and instead focus on the acoustic differences that were in fact important for identifying the L2 sounds.

Lively, Logan, and Pisoni then recreated their study in 1993, this time with the addition of a generalization task, in which participants needed to identify new words produced by familiar and unfamiliar voices. Participants were divided into two groups - one group trained with stimuli from a variety of voices, and the other group trained with stimuli from only one voice. The researchers found that the variability was an important factor in the category formation of L2 sounds. A third study by Lively, Pisoni, Yamada, Tohkura, and Yamada (1994), revealed that this high variability phonetic training was also successful for retaining learning long term, after testing with 3-month and 6-month delayed post-tests.

Two types of phonetic training, forced-choice identification training and discrimination training, will be discussed in the following section.

2.3 FORCED-CHOICE IDENTIFICATION VERSUS DISCRIMINATION

Two types of perception training that have been researched for L2 perception are forced-choice identification training and discrimination training. Identification training refers to training in which listeners are presented with one stimulus at a time and need to label the sounds into a set of defined categories (PISONI, 1971; POLLACK.) There are several types of discrimination training, including ABX, when listeners are presented with three stimuli and they need to choose whether the third sound is more similar to the first or second sound. (POLLOCK et. al, 1971.) Another type of discrimination training is with the 4IAX test, in which listeners are presented with four sounds and need to select whether the first or second pair is more alike (POLLOCK et. al, 1971.) The 2IAX test is yet another discrimination test, where listeners are presented with two sounds and need to choose whether the sounds are similar or different. (POLLOCK et. al, 1971.)

A study with forty-five Spanish/Catalan bilinguals in their first year of a Spanish university English program introduced six thirty-minute sessions of perception training for English vowels to participants (CEBRIAN; CARLET; GORBA; GAVALDA, 2019). Specifically, the intervention was high variability perceptual training, and utilized both identification and discrimination training. The results of this study showed that training was in fact effective in improving the participants' identification and discrimination of L2 vowels. The post-test (conducted four months later) scores also revealed long-term effects of the perception training.

Aliaga-García also conducted a study with advanced Spanish/Catalan bilingual learners of English, investigating the effects of phonetic training sessions, specifically including perception training with forced-choice identification tasks and discrimination tasks (ALIAGA-GARCÍA, 2009). Participants in this study completed six two-hour training sessions. The results did not reveal overall significant gains in perceptual competence for all the sound pairs in the study; however, the perception training did appear to significantly improve learners' discrimination ability for the vowels /i ɪ æ ʌ/ (ALIAGA-GARCÍA, 2009).

Another study compared the effects of identification versus discrimination training for both L2 perception and production, this time working with adult Japanese learners of English and the English consonant sounds, /r/ and /l/ (SHINOHARA; IVERSON, 2018). In this study, participants were given ten sessions of identification and discrimination training, with pre-, mid-, and post-tests of identification, auditory discrimination, category discrimination, and production (SHINOHARA ET. AL, 2018). The results revealed that both training methods were effective in increasing participants' perception and production; however, there was not a significant benefit to using the two training methods together in combination (SHINOHARA ET. AL, 2018).

One study looked at the effects of these two types of training for one-hundred bilingual Spanish/Catalan learners of English and their perception of English vowels and initial and final stops (CARLET; CEBRIAN, 2015). The vowels that were trained and tested in this study were /i ɪ æ ʌ ɜ/ (CARLET, 2015). This study revealed that for L2 stops, both training methods (ID – identification; DIS – discrimination) were equally effective; however, for vowels, the "ID trainees improved and generalized learning to greater extent than the DIS trainees on the perception of L2 vowels" (CARLET, 2015, p. 944). Carlet provided a possible explanation for the superiority of the ID training method versus the DIS training method for promoting generalization, noting that "this superiority might be connected to the presence of labels in the ID task, which provided learners with focus on phonetic form (i.e., phonetic symbols and/or orthography), which is said to impact speech perception" (CARLET, 2015, p. 946).

Nozawa also compared the effects of these two types of training, working with native Japanese learners of American English (NOZAWA, 2015). In this study, listeners were divided into four groups, two groups vowel-oriented and two groups nasal-oriented (NOZAWA, 2015). One of the vowel-oriented groups received vowel identification training, and the other received vowel discrimination training (NOZAWA, 2015). One of the nasal-oriented groups received nasal identification training and the other received nasal discrimination training (NOZAWA, 2015). The results of this study revealed that training did not have an effect on nasal identification; however, it did have an effect on vowel identification (NOZAWA, 2015). The vowel-oriented group with identification training showed the most gains compared to the other three groups (NOZAWA, 2015).

It is also important to note that some studies have found little difference between the effects of forced-choice identification and categorical discrimination training for L2 perception. For example, Flege found in his 1995 study that native speakers of Mandarin who learn English as a second language did not show a significant difference for identification versus discrimination training in their perception of /t/ and /d/ in the final position of English words (FLEGE, 1995).

Considering the results of the studies mentioned in this section and the indications that perhaps there is not a significant benefit to using both training methods (SHINOHARA ET.

AL, 2018) and that between the two methods, identification training may be more effective (CARLET, 2015; NOZAWA, 2015), the present research will utilize identification perception training.

2.4 NATURAL VERSUS SYNTHETIC STIMULI FOR PERCEPTION TRAINING

Perception training with synthetic stimuli "has been found to be effective in changing L2 cue weighting of non-native vowel perception" (CHENG ET. AL, 2019, p.4). An acoustic cue is defined as "any acoustic characteristic of a segment which aids in the recognition of that segment in speech" (TRASK, 2006), and cue weighting refers to the importance listeners give to certain acoustic cues to distinguish and recognize specific sounds. Cue weighting will be especially important for the present study, as the synthetic stimuli will aim to assist learners in focusing on the formant frequencies as cues. Native English speakers utilize both durational and formant frequencies as cues for identifying vowel sounds, although they rely primarily on the latter (ESCUDERO ET AL., 2009). These formant frequency cues may be more easily generalized to new contexts.

The following paragraphs will discuss studies by Cheng et. al, Nobre-Oliveira, Ylinen et. al, and Wang et. al, which have indicated the effectiveness of synthetic stimuli for perception training, specifically to help learners adjust their cue weighting and improve their ability to distinguish vowel pairs in various contexts.

Cheng et. al, for example, found that synthetic stimuli were in fact useful in assisting learners to focus on the most important cues and essentially ignore the less important features (2019). Cheng's study focused specifically on the role of temporal acoustic exaggeration in perception training for Chinese adults learning the English /i- ɪ/ vowel contrast. Two groups of participants received perception training for the /i- ɪ/ vowel pair, one group listening to naturally produced words by American English speakers and one group listening to acoustically modified words (the vowel duration adjusted to 170 ms) by the same American English speakers (CHENG ET. AL, 2019). Both groups in this study showed significant improvement in their ability to distinguish these two vowels, although the group that listened to acoustically modified stimuli was more successful in generalizing their learning to new speakers and words (CHENG ET. AL, 2019).

Nobre-Oliveira researched the effect of synthetic stimuli on intermediate-level native BP speakers of L2 English. She manipulated and enhanced F1 and F2 values of the vowel pairs

/i-ɪ/, /æ-ɛ/, and /ʊ-u/ in synthetic stimuli in order to create a larger vowel space, providing an easier distinction for learners to categorize sounds (NOBRE-OLIVEIRA, 2007). In order to further enhance the stimuli, this study used isolated vowels (versus vowels within words), which allowed for the control of vowel duration (NOBRE-OLIVEIRA, 2007). The duration of all vowels for perception training was 500ms, which is longer than typical vowel production in real speech (NOBRE-OLIVEIRA, 2007). However, this was done intentionally in order to enhance the "different spectral properties of each vowel during the training phase, which would hopefully help learners to improve their ability of categorizing L2 vowels successfully" (NOBRE-OLIVEIRA, 2007, p. 88). The final results of this study revealed that synthetic stimuli are at least equally, if not more, effective as natural stimuli in perception training (NOBRE-OLIVEIRA, 2007). However, this data was not statistically significant (NOBRE-OLIVEIRA, 2007).

Ylinen et. al found that perception training resulted in plastic changes in the early processing stages in the cortex for the weighting of phonetic cues for Finnish second-language users of English (YLINEN ET.AL, 2009). This study worked with ten native Finnish speakers and investigated the use of spectral and duration cues for recognizing the English /i-ɪ/ vowels. The pre-training tests revealed that Finns did in fact rely more on duration rather than formant frequencies for vowel recognition (YLINEN ET. AL, 2009), in accordance with Escudero et. al's suggestion that L2 learners of English tend to rely more heavily on durational rather than formant frequency cues (2004). However, perception training in this study was successful in assisting Finnish users of English to use spectral cues more reliably in the recognition of English words with the /i-ɪ/ vowels, and it appears that the use of stimuli with controlled duration was a key factor in the cortex-level changes in cue weighting for participants (YLINEN ET. AL, 2009).

Wang et. al. found that fourteen adult native speakers of Mandarin were able to effectively shift their attention from temporal properties to spectral properties in English vowel perception tasks following a short period of training with synthetic stimuli (WANG ET. AL, 1999). The stimuli used for the perception training were /hid/, /hɪd/, /hud/, and /hʊd/ (WANG ET. AL, 1999). Participants for this study participated in perception training sessions with immediate feedback, which were held for each participant repeatedly (up to a maximum of four training sessions) until a minimum score of 95% accuracy was achieved (WANG ET. AL, 1999). These four tokens were synthesized with six temporal and six spectral steps, generating a total of 36 tokens (WANG ET. AL, 1999). This study revealed Mandarin listeners' tendency to rely heavily on the temporal cues for the contrast between the English vowels /i-ɪ/, and a lack of a sufficiently strong response to spectral cues in both /i-ɪ/ and /ʊ-u/ (WANG ET. AL, 1999). The results from this study showed that even a short period of training with feedback may be very effective in assisting L2 learners to shift their attention from temporal to spectral properties in vowel perception tasks (WANG ET. AL, 1999).

2.5 BP AND ENGLISH VOWEL INVENTORIES AND DISTINCTIONS

The General American English vowel inventory includes around 14-15 different vowels (LADEFOGED, 2007) and the Brazilian Portuguese vowel inventory includes 7 oral vowels and 5 nasal vowels (SEARA; NUNES; LAZZAROTTO-VOLCÃO, 2015). Of these vowel inventories, there are some similarities, one of those being that they both include the vowel /ɛ/.

Vowels are categorized acoustically based on formant frequencies, amplitude, and duration (SILVA ET. AL, 2019). The major formant frequencies are referred to as F1, F2, and F3. F1 notes the height, or vertical dislocation (of the tongue), F2 refers to backness, or horizontal dislocation (of the tongue), and F3 refers to the roundness (of the lips) (SILVA; SEARA; SILVA; RAUBER; CANTONI, 2019). The F4 and F5 formants provide additional information about the vowel quality (SILVA ET. AL, 2019). Silva et. al note that the F1 and F2 formants are the most important ones for vowel identification (2019).

Regarding the Brazilian Portuguese vowel /ɛ/, the average F1 values are 646 for females and 518 for males. (ESCUDERO; BOERSMA; SCHURT RABUER; BION, 2009). The average F2 values are 2,271 for females and 1,831 for males (ESCUDERO ET. AL, 2009). The average F3 values are 2,897 for females and 2,772 for males (ESCUDERO ET. AL, 2009). The duration of the Brazilian Portuguese vowel /ɛ/ is approximately 141 ms for females and 123 ms for males (ESCUDERO ET. AL, 2009).

The General American English vowel /ɛ/ shows relatively similar characteristics as the Brazilian Portuguese /ɛ/, although they are not exactly the same. The average F1 values for the General American English /ɛ/ are 731 for females and 580 for males (HILLENBRAND; GETTY; CLARK; WHEELER, 1995). The average F2 values are 2,058 for females and 1,799 for males (HLLENBRAND ET. AL, 1995). The average F3 values are 2,979 for females and 2,605 for males (HILLENBRAND ET. AL, 1995). The average duration for the English /ɛ/ is 254 ms for females and 189 ms for males (HILLENBRAND ET. AL, 1995). Compared to the

Brazilian Portuguese /ɛ/ therefore, the formant frequencies appear to be similar, whereas the duration is notably shorter than the General American English /ɛ/.

The vowel /æ/ is not a part of the Brazilian Portuguese vowel inventory. In General American English, this vowel has average F1 values of 669 for females and 588 for males (HILLENBRAND ET. AL, 1995). The F2 values are approximately 2,349 for females and 1,952 for males (HILLENBRAND ET. AL, 1995.) The average F3 values are 2,972 for females and 2,601 for males (HILLENBRAND ET. AL, 1995). The average duration is 332 ms for females and 278 ms for males (HILLENBRAND ET. AL, 1995).

While comparing the English vowels /æ-ɛ/, it can be noted that according to Hillenbrand et. al's data, the major distinction between these vowels lies in the F2 values and the duration (HILLENBRAND ET.AL, 1995). The F1 and F3 values are very similar – the F1 values differ only by 62 for females and 8 for males, and the F3 values differ only by 7 for females and 4 for males (HILLENBRAND ET. AL, 1995). However, the Brazilian Portuguese /ɛ/ average F2 values fall right between the average F2 values of the English /æ/ and /ɛ/, likely causing a difficulty for BP learners of English for distinguishing these two vowels. Research has shown that distinguishing these two vowels is in fact especially difficult for BP learners (LIMA JR., 2017; RAUBER, 2006), as will be discussed in the following section.

2.6 BP LEARNERS' ACQUISITION OF THE ENGLISH VOWELS /Æ-Ɛ/

As discussed previously, SLM predicts that L2 learners will often struggle to form categories for L2 sounds that are similar to their L1 sounds, and as a result assimilate the two, creating a merged L1-L2 (FLEGE, 1995).

Lima Jr. (2017) notes that the vowel pairs /i-ɪ/, /æ-ɛ/, and /ʊ-u/ are some of the most difficult for BP learners of English to perceive, as these learners typically assimilate the pairs into the BP categories pairs /i/, /ɛ/, and /u/, respectively. He also found a hierarchy of difficulty with these vowel pairs, with /i-ɪ/ being the easiest to perceive, followed by /u-ʊ/, and finally /æ-ɛ/ being the most difficult (LIMA JR., 2017). Rauber (2006) also found that the vowel pair /æ-ɛ/ was the most poorly perceived pair (compared to /i-ɪ/ and /u-ʊ/) for BP learners of English. This study will focus on the vowel pair /æ-ɛ/ because of this tendency for it to pose the biggest difficulty for BP learners of English.

Following previous studies (CHENG et. al, 2019; NOBRE-OLIVEIRA, 2007) the researcher chose to compare the use of natural versus synthetic stimuli for perception training in order to better understand whether one may be more effective than the other for BP L2 English learners. The researcher specifically decided to manipulate the duration of the vowel sounds to help learners adapt their cue weighting and depend more consistently on formant frequency differences to identify the vowel sounds. Identification training (instead of discrimination training) was selected because of previous research's indication that it may be more effective (CARLET, 2015; NOZAWA, 2015). In the next chapter, the method used for instrument preparation, data collection, and data analysis will be discussed.

3 METHOD

3.1 INTRODUCTION

The main purpose of this study is to investigate the effectiveness of synthetic versus natural stimuli in perception training for Brazilian learners of English. The study also aims at identifying whether L2 learners are able to generalize their learning to new voices and stimuli, determining the delayed effects of training, and understanding the relationship between proficiency level and the effectiveness of perception training. This chapter presents the methodological design of the present study, including the preparation of stimuli for the pre-test, training, and post-tests, the participants for the study, the instruments used, and the procedures for data collection and analysis.

Section 3.2 describes the participants and the three groups into which they were divided. Section 3.3 discusses the instruments for data collection and specifically the process of stimuli preparation. In sections 3.4 and 3.5, the procedures for data collection and data analysis are explained.

The data collection was carried out online in May 2021. The study's proposal translated into Portuguese, a participant consent form, and all of the research instruments were submitted to and approved by the *Comitê de Ética em Pesquisa com Seres Humanos da UFSC*[2].

3.2 PARTICIPANTS

Fifty-six native Brazilian-Portuguese speakers and learners of English participated in this study. These participants were invited to participate through the social media platform, Instagram. The participants ranged in age from nineteen to forty-nine, and they ranged in proficiency from beginner to proficient user of English as an L2. They did not have any hearing or vision impairment at the time of the study.

Participants did not receive financial compensation for their participation in the study, per the guidelines set in place by the Brazilian Ethics Committee. The identity of all participants will remain anonymous, and they will be referred to as P1, P2, and so on in the study's report.

---

[2] Parecer Number: 4.622.381

The participants were divided randomly into three groups: 1) those that received perception training with natural stimuli, 2) those that received perception training with synthetic stimuli, and 3) those that did not receive any perception training and served as a control group. There were 22 participants in the natural group, 15 participants in the synthetic group, and 19 participants in the control group.

3.3 INSTRUMENTS FOR DATA COLLECTION

The instruments for data collection in this study include an online background questionnaire, an online vocabulary test, a perception identification pre-test, post-test, and delayed post-test, and perception identification training activities. The online background questionnaire was filled out by participants using Google Forms (www.google.com/forms). The vocabulary test used was the Lognostics online Yes/No test (www.lognostics.co.uk/tools/V_YesNo/V_YesNo.htm). The perception tests and the perception training were presented to the participants with the help of Gorilla Experiment Builder (www.gorilla.sc.)

The natural and synthetic stimuli were reviewed by native English and Brazilian proficient users of English as an L2 and underwent pilot testing with a limited number of participants prior to data collection. The following sections will describe in more detail the questionnaire as well as the vocabulary test, perception pre-test, perception training, and the perception post-tests that were conducted using Gorilla Experiment Builder.

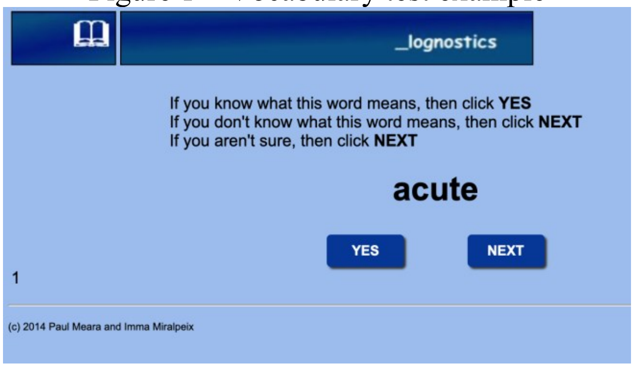
3.3.1 Background questionnaire

An online background questionnaire on Google Forms was used to collect background information about participants' L2 learning experience, current L2 self-reported level, and knowledge of other languages. The questionnaire also asked participants to report any hearing or vision problems. The main purpose of this questionnaire was to identify factors that may cause individual differences in the perception identification pre-test and post-tests scores (such as knowledge of another language that discriminates between the /æ-ɛ/ vowels). The full questionnaire can be found in Appendix B.

3.3.2 Vocabulary test

The Lognostics online Yes/No test[3] was carried out with the purpose of evaluating participants' proficiency level in English. This test uses the Yes/No methodology that was developed by Meara and Jones in the Eurocentres Vocabulary Size Test (1989). As shown in Figure 1, individual vocabulary words were shown on screen and participants were asked to select "Yes" if they knew the meaning of the word and "Next" if they were unsure or didn't know the meaning of the word. There were a total of 200 words presented to each participant. The test took a total of about 10 minutes to complete. Participants were able to keep track of how many items they had responded to, as the number was recorded in the bottom left corner, as seen in Figure 1.

Figure 1 – Vocabulary test example



Source: Lognostics (MEARA, MIRALPEIX, 2014)

Once completed, the participants were shown a final report screen with a total score, as illustrated in Figure 2. The maximum score for this test was 10,000. Based on the score descriptions provided in the test manual (MEARA; MIRALPEIX, 2014), scores between 2,000-3,500 are typical for competent beginners, scores between 3,500-6,000 are typical for intermediate learners, and scores between 6,000-10,000 can be considered an indication of advanced proficiency level.

[3] www.lognostics.co.uk/tools/V_YesNo/V_YesNo.htm

Figure 2 – L2 English level test report screen

Source – Lognostics (MEARA, MIRALPEIX), 2014.

3.3.3 Stimuli preparation

This section will describe the process for stimuli preparation. Section 3.3.4 includes detailed information about the selection of speakers for the voice recordings. Section 3.3.4.1 discusses the carrier words that were used in the study, and section 3.3.4.2 describes the software and processes that were used in order to edit the recordings to create the stimuli for this study.

3.3.4 Speaker selection for stimuli

Traditionally, research in the area of perception training has used native speakers' voices for stimuli creation; however, the present study will utilize a mixture of both native (American) speakers and non-native (Brazilian) proficient speakers. English as a Lingua Franca (ELF) is now widely used by speakers of various L1s, and a common context for communication is between two non-native speakers (JENKINS, 2009). With this in mind, ideally the stimuli would be created from a wider variety of speakers (including both ELF speakers with various L1s and native speakers from various geographical regions), however due to the time constraint of a master's research project, the researcher selected two speaker backgrounds that were believed to be most relevant to the participants and their current learning context as well as those that were most accessible to the researcher at the time of stimuli preparation – native American English speakers (which participants may encounter in movies, music, and YouTube videos) and proficient Brazilian L2 speakers (which participants might encounter in-person in Brazil, in the media, in the classroom, or online).

The researcher recruited a total of thirteen speakers (four native American English speakers[4] and eight native BP speakers of English) who sent in audio files of voice recordings, which were then edited using PRAAT and Audacity to cut the files, remove background noise, and normalize amplitude (this process will be described in more detail in section 3.3.3.2.). The researcher then presented the recordings from these thirteen speakers to three listeners (one native American English speaker and two native BP speakers of English) for review. These listeners provided detailed feedback on each individual audio, noting which carrier word they identified, the ease or lack thereof that they had to identify it, and the quality of the audio file. Based on the feedback from these three listeners, the researcher then selected recordings from seven speakers (four native American English speakers and three native BP speakers of English), which can be seen in Table 1. The identity of all speakers will remain anonymous, and they will be referred to as S1, S2, and so on in the study's report.

Table 1 – English speakers for stimuli creation

Speaker	Nationality	Gender	Age	Education	English
S1	American - Washington state	Female	58	Some college	Native speaker
S2	American - Washington state	Male	20	Some college	Native speaker
S3	Brazilian – Pará	Female	35	Master's student	Proficient user of English as a L2
S4	Brazilian – Bahia	Male	35	Doctorate student	Proficient user of English as a L2
S5	American - Washington state	Male	66	Undergradu ate degree	Native speaker
S6	American - Washington state	Female	25	Graduate student	Native speaker
S7	Brazilian - Santa Catarina	Male	31	Undergradu ate degree	Proficient user of English as a L2

SOURCE: The author (2020).

4 In order to maintain consistency and to facilitate stimuli collection, the research chose four native American English speakers all from the same state, Washington.

After the final preparation of the stimuli and perception identification tests and training, the researcher presented the stimuli from these seven selected speakers a second time to native American English and native BP speakers of English for review as part of a pilot study. In this pilot, the stimuli were correctly identified by two native American English speakers 100% of the time, and by two native BP speakers of English 88.1% of the time. These listeners who reviewed the stimuli have been described in more detail in Table 2.

Table 2 – Listeners for stimuli review

Listener	Nationality	Gender	Age	Education	English
LN1	American - Washington state	Female	28	Associate's degree	Native speaker
LN2	American - Oregon state	Female	23	Bachelor's degree	Native speaker
LN3	Brazilian - Bahia	Female	34	Doctorate student	Proficient user of English as a L2
LN4	Brazilian - Santa Catarina	Male	25	Bachelor's degree	Proficient user of English as a L2

Source: The author (2020).

In addition to review by native American English and native BP speakers of English, the vowel duration, F1, F2, and F3 values of the stimuli have been measured using PRAAT and saved in spreadsheets (Appendix C). As can be seen in the appendix, recordings from both the native American English speakers and the native BP speakers of English varied slightly in vowel duration and formant frequencies. However, all stimuli maintained a coefficient of variation of 0.26 or less for vowel duration, F1, F2, and F3 values, except for the female recordings of the distractor vowel /i/, which had a coefficient of variation of 0.40 for vowel duration.[5] These coefficients of variation values gave the researcher confidence[6] that the

[5] The mean, standard variation, and coefficient of variation values, as well as the individual vowel duration and formant frequency values for all stimuli can be found in Appendix C.

[6] Banik, Kibria, and Sharma note that a coefficient of variation less than 0.33 is considered very accurate, and a coefficient of variation between 0.33 and 0.67 is considered reasonably accurate (2012). For this reason, the author was confident with the coefficient of variations less than or equal to 0.26.

recordings from both the native American English speakers and from the native BP speakers of English were appropriate for use as stimuli in this study.

3.3.4.1 Carrier words

The stimuli for this study are composed of twenty-one single-syllable carrier words, divided into seven triads. Following the use of distractors in previous research on perception training (NOBRE-OLIVEIRA, 2007), the researcher decided to include a distractor within each triad. Each of these triads included one word with the target-vowel /æ/, one word with the target-vowel /ɛ/, and one word to be used as a distractor with the vowel /i/. All tokens were one-syllable words, and the vowels appeared in each token following a single onset phoneme and followed by a single stop consonant.

These carrier words were recorded by the seven English speakers listed in Table 1. Each speaker recorded one triad, a total of three carrier words. These carrier words can be seen below in Table 3.

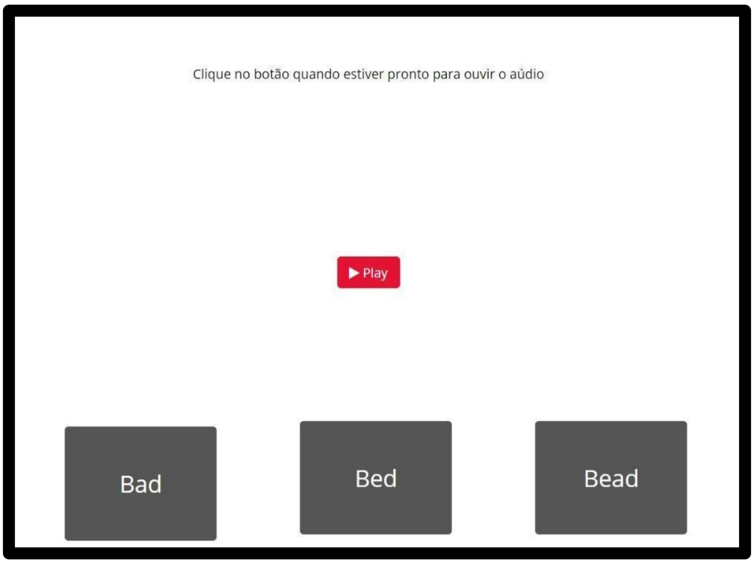
Table 3 – Carrier words recorded by seven adult speakers

Type of stimuli	Vowel	Carrier words						
		S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**
Target	**/æ/**	bad	had	mat	bat	vat	sat	sad
	/ɛ/	bed	head	met	bet	vet	set	said
Distractor	**/i/**	bead	heed	meet	beat	Veet	seat	seed

Source: The author (2020).

In effort to control for the phonemic environments of the target vowels, the researcher used some low-frequency vocabulary as carrier words (notably, “heed”, “vat”, and “Veet”). Because this study involved L2 English students of various proficiency levels, the researcher acknowledged that these low-frequency words, as well as differences in orthography (for example, words like “said” and “head,” which have the same vowel sound but are represented by different spellings,) could interfere with participants’ ability to correctly identify words. In effort to lower this effect and make the perception identification tests and training activities easier (especially for the beginner level students), the carrier words within each triad were always presented to participants in the same order on screen: the carrier word with the target vowel /æ/ on the left, the carrier word with the target vowel /ɛ/ in the middle, and the carrier word with the distractor vowel /i/ on the right (for example, "Bad" on the left, "Bed" in the middle, and "Bead" on the right). An example of this can be seen in Figure 3 below.

Figure 3 – Carrier word presentation



SOURCE: The author (2021)

3.3.4.2 Stimuli alterations with PRAAT and Audacity

In order to prepare the stimuli, the researcher first instructed the seven speakers to send audio files recorded on their cellphones to the researcher[7]. The speakers recorded each word in a separate audio, repeating the same word three times. The researcher then transferred the audios to the computer and saved them as .wav files.

The files were then edited using the software Audacity to remove background noise and to normalize the amplitudes, the peak amplitudes of all audio files being between 0.900 and 1.001 dB. A set of procedures were taken in order to remove background noise using the Audacity audio editor software. First, the researcher selected a sample of the kind of noise to be removed from the file. She then selected "effects," then "remove noise," then "obtain noise profile." Next, the researcher selected the whole file and then clicked on "effects" and then "remove noise." The researcher then previewed the file for audio quality and finally saved the

[7] Because the native American English speakers were in the United States during the time of stimuli collection and preparation, the researcher asked these speakers to send cellphone recordings (instead of conducting in-person recordings). To maintain consistency, the researcher decided to use the same method for the recording collection of the Brazilian L2 English speakers, who were in various parts of Brazil at the time of collection.

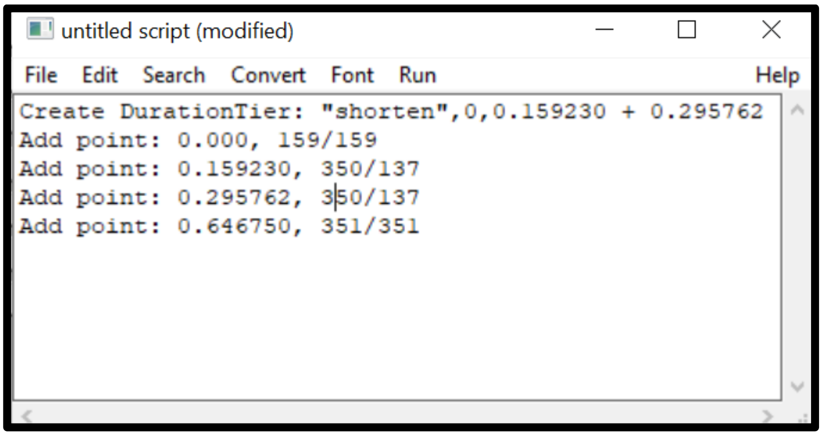
new file, exporting it as a .wave file. In order to normalize the amplitudes for the audio files, the researcher first selected “tools,” then “macro,” then “new.” Next, she typed “normalize” in the box and selected “ok.” Next, the researcher selected “apply macro to” and selected all of the audio files that were to be normalized. Finally, she inspected the sound waves of the normalized files to verify that they all had similar amplitudes.

The files were also edited using PRAAT, to cut the audio files to include only one repetition of each carrier word and then to adjust the target vowel duration of the synthetic stimuli to 350ms. In order to adjust the target vowel duration, the following procedures were taken, following instruction from the PRAAT software itself. First, the researcher selected the natural cropped audio file already saved in PRAAT. Next, she selected “manipulate” and then “to manipulation.” At this point, the time step, minimum pitch, and maximum pitch were not altered. The researcher then selected the new manipulation audio listed in the dynamic menu and clicked on “Praat” and then “New Praat script.” The researcher then wrote the following new Praat script into the script window:

> Create DurationTier: "shorten", [beginning of vowel sound] + [end of vowel sound]
> Add point: 0.000, [duration of initial consonant sound] / [duration of initial consonant sound]
> Add point: [beginning of vowel sound], 350/[duration of original vowel sound]
> Add point: [end of vowel sound], 350/[duration of original vowel sound]
> Add point: [end of final consonant sound], [duration of final consonant sound] / [duration of final consonant sound]

After writing in the script, the researcher then went to the Praat window and selected “run” and then “run” again. On the dynamic menu, the researcher then selected both the manipulation object and the new duration tier. She then selected “Replace duration tier.” Next, she selected the manipulation object and then clicked “Get resynthesis (get overlap-add).” She then selected “rename” and renamed the file as “[carrier word] (synthetic) – Speaker (#)”. Finally, the file was saved as a wav file.

Figure 4 – Example of PRAAT script for the alteration of the vowel duration in the carrier word “sat”



Source: The author (2020).

The vowel duration of the synthetic stimuli was manipulated to be 350 ms, purposefully longer than the average vowel production in natural speech with the intention of providing more time for the learners to note formant frequency distinctions, as Nobre-Oliveira did in her 2007 study. Although Nobre-Oliveira selected 500ms as the durational value in her study (2007), the researcher for the present study selected 350ms, as it is still longer than the average vowel production of the Brazilian Portuguese /ɛ/[8], General American English /ɛ/[9], and General American English /æ/[10] but allows for higher audio quality as it involves less alteration to the original sound.

3.3.5 Perception identification pre-test

The stimuli used in the pre-test included recordings from Speakers 1, 2, 3, and 4[11] and can be noted in Table 4. The stimuli for the pre-test included a total of twelve tokens, which were organized into triads. The duration of these carrier words was not altered for the pre-test.

---

[8] The duration of the Brazilian Portuguese vowel /ɛ/ is approximately 141 ms for females and 123 ms for males (ESCUDERO ET. AL, 2009).

[9] The average duration for the General American English /ɛ/ is 254ms for females and 189ms for males (HILLENBRAND ET. AL, 1995).

[10] The average duration for the General American English /æ/ is 332ms for females and 278ms for males (HILLENBRAND ET. AL, 1995).

[11] The researcher attempted to evenly distribute the speakers between those whose recordings were used for the pre-test, training, and post-tests, and those whose recordings were used as unfamiliar stimuli only in the post-tests.

Table 4 – English stimuli for pre-test and training

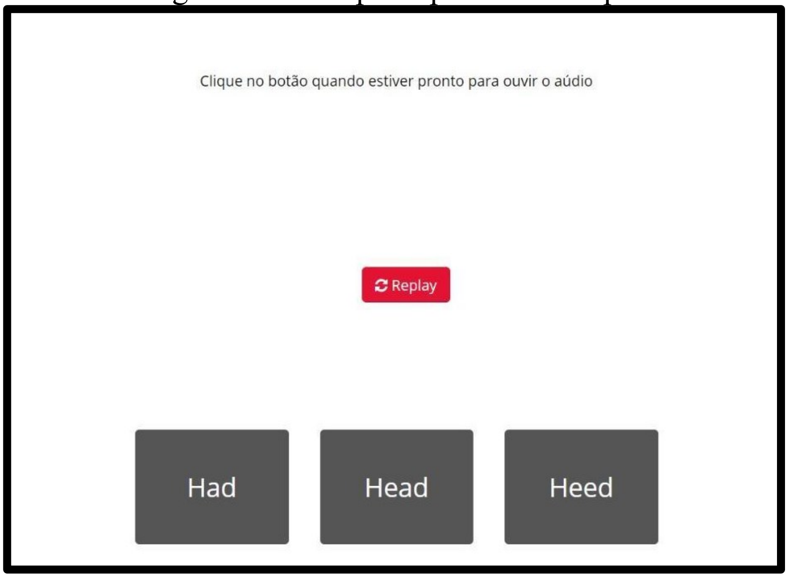
Type of Stimuli	Vowe l	One-syllable token words			
		S1	**S2**	**S3**	**S4**
Target	**/æ/**	bad	had	mat	bat
	/ɛ/	bed	head	met	bet
Distractor	**/i/**	bead	heed	meet	beat

SOURCE: The author (2020).

These carrier words were presented to participants using the online research platform Gorilla Experiment Builder. Participants were shown each of the four triads listed in Table 4 a total of six times (twice for each carrier word). An example of a triad presentation is shown in Figure 5. Each time that a triad appeared on screen, the participants clicked the "play" button when they were ready. They would then hear the carrier word and need to select which word they heard, selecting a button on the screen accordingly. Each of these carrier words was played twice (for a total of six audios per slide), and participants were allowed to listen to the same audio up to two times if needed. Figure 5 shows a screenshot of the perception identification pre-test, designed with Gorilla Experiment Builder. The figure contains a triad with the three response options, the button "Play" which the participant selects when he or she is ready to listen to the stimulus, and the button "Replay" (appears after the participant plays the audio the first time) in case the participant would like to repeat the audio.

The perception identification pre-test and training utilized recordings from one male native English speaker, one female native English speaker, one male Brazilian L2 English speaker, and one female Brazilian L2 English speaker. The unfamiliar stimuli to be used only in the post-tests were recorded by one male native English speaker, one female native English speaker, and one male Brazilian L2 English speaker.

Figure 5 – Perception pre-test example



Source: The author (2021).

The stimuli for the pre-test were automatically randomized for the participants via the platform Gorilla Experiment Builder. This was done following previous research (NOBRE-OLIVEIRA, 2007) in order to prevent participants from memorizing responses.

In total, the perception pre-test included twenty-four opportunities for participants to respond (two repetitions of each carrier word within each triad).

3.3.6 Perception identification training

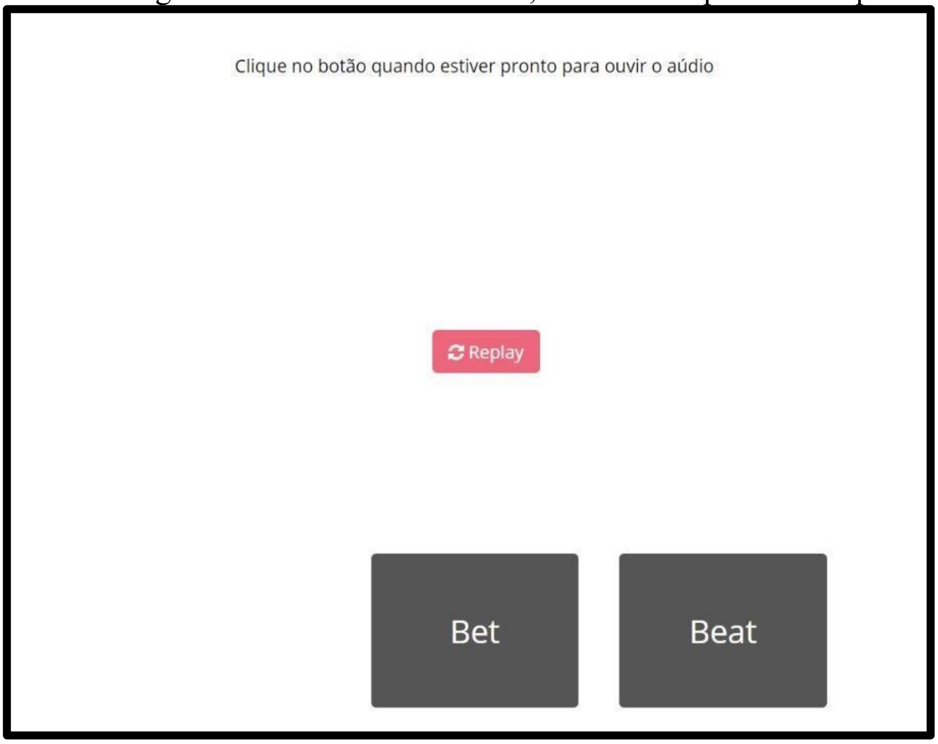
The participants assigned to the control group did not receive perception training. The participants assigned to the natural stimuli experimental group received three perception training sessions with recordings that did not have their duration altered. The participants assigned to the synthetic stimuli experimental group received four perception training sessions with recordings that had their duration altered to 350ms. The platform Gorilla Experiment Builder randomly sorted participants into these three groups.

These three training sessions for participants in the two experimental groups utilized the same stimuli from the perception identification pre-test, as seen in Table 4, however, these carrier words were presented to participants either naturally or synthetically, depending on the experimental group to which they were assigned.

All the students participating in the perception training listened to a total of twenty-four (either natural or synthetic, depending on their experimental group) stimuli, presented in triads as shown in Figure 5 on Gorilla. The training sessions lasted approximately ten minutes.

For each triad, participants heard a stimulus and selected which of the three alternatives displayed on the screen they heard. Each stimulus was presented two times per set, and participants were able to listen to the stimulus up to two times if necessary. Participants received immediate feedback following each response. When they responded correctly, they automatically moved on to the next triad. When they responded incorrectly, their incorrect answer disappeared from the screen, leaving two options (the remaining two carrier words) for them to choose from for a second attempt (as seen in Figure 6). These three training sessions were conducted over the course of three consecutive days.

Figure 6 – Immediate feedback, incorrect response example



Source – The author (2021).

As the duration of both the /æ/ and /ɛ/ (and /i/) vowels were the same for the synthetic stimuli, 350 ms, participants were unable to depend on duration for vowel identification. The purpose of this adjustment to the stimuli was to assist learners in their cue weighting and depend more consistently on the formant frequency values for vowel identification. Controlling duration was found to successfully change L2 cue weighting for Chinese learners of English's perception of the English vowels /i-ɪ/ (CHENG ET. AL, 2019), Finnish users of English's perception of the English vowels /i-ɪ/ (YLINEN ET. AL, 2009), and native speakers of Mandarin's perception of the English vowels /i-ɪ/ and /ʊ-u/ (WANG ET. AL, 1999). According to Hillenbrand's 1995 data, the F2 value differences are the most significant (compared to the other formant values) for vowel distinction. These values should be naturally emphasized as the duration is controlled.

3.3.7 Perception identification post-test

As seen below in Table 5, the post-test was composed of a combination of familiar tokens (the same recordings from speakers 1, 2, 3, and 4 from the pre-test and training sessions) and new tokens (recordings from speakers 5, 6, and 7 that have not been previously introduced to the participants). All these stimuli were presented naturally, without alteration to the vowel duration. The perception post-test was conducted in the same format as the perception pre-test as shown in Figure 5 and Figure 6, but due to the higher amount of stimuli being presented, took approximately 20 minutes (instead of 10 minutes).

Table 5 – Perception post-test stimuli triads

Type of stimuli	Vowel	Carrier words						
		Familiar speakers and carrier words				Unfamiliar speakers and carrier words		
		S1	S2	S3	S4	S5	S6	S7
Target	/æ/	bad	had	mat	bat	vat	sat	sad
	/ɛ/	bed	head	met	bet	vet	set	said
Distractor	/i/	bead	heed	meet	beat	Veet	seat	seed

Source: The author (2020).

The stimuli for the perception post-test were automatically randomized by the research platform Gorilla Experiment Builder.

3.3.8 Delayed perception identification post-test

The delayed perception post-test was the exact same format as the perception post-test; however, it was presented to participants six days following the completion of the post-test. The stimuli used for this delayed post-test can be seen in Table 5 above. This delayed post-test took approximately 20 minutes for participants to complete.

3.4 PROCEDURES FOR DATA COLLECTION

This section will describe in detail the procedures for data collection. Section 3.4.1 will discuss the online background questionnaire that participants completed prior to the perception identification pre-test. Section 3.4.2 will describe the vocabulary test that participants took prior to the perception tests, and Section 3.4.3 will explain how the data from the perception identification pre-test, post-test, and delayed post-test was collected. See Table 6 for reference of the time frame for each stage of the data collection.

Table 6 – Data collection schedule

Recruitment	03/05/2021 - 06/05/2021
Consent Form and Background Questionnaire	04/05/2021 - 06/05/2021
English Level Vocabulary Test and Perception Pre-Test	07/05/2021 - 08/05/2021
Perception Training #1	11/05/2021
Perception Training #2	12/05/2021
Perception Training #3	13/05/2021
Perception Post-Test	14/05/2021
Delayed Perception Post-Test	20/05/2021

Source - The author (2021).

3.4.1 Online questionnaire data collection

After the study was approved by the *Comitê de Ética em Pesquisa com Seres Humanos da UFSC*, the researcher published social media announcements inviting Brazilian learners of English to participate in the study. Participants that were interested completed the online background questionnaire via Google Forms and a consent form. This background questionnaire detailed their experiences learning English, knowledge of other languages, and noted any hearing or vision problems.[12]

3.4.2 English level assessment

Participants that were eligible then scheduled group Zoom video call meetings with the researcher. During the first encounter, the participants completed an L2 English level assessment (as well as the identification perception pre-test.) The level test was conducted using the Lognostics Yes No Vocabulary test by Paul Meara and Imma Miralpeix. Participants received the web link to begin the test through Zoom. The description of this test can be found in Section 3.3.2. Once the participants completed the test, they reported their scores (shown on the final report screen on the test program) to the researcher during the video call, and the researcher manually recorded them.

3.4.3 Perception tests data collection

During the same Zoom meeting, participants then took a perception pre-test to evaluate their ability to identify the English vowels /æ-ɛ/. After reporting their vocabulary test score, the researcher sent a second web link to the participants through Zoom, which they used to access the pre-test. The participants logged into Gorilla Experiment Builder using a participant ID number pre-assigned by the researcher. The perception pre-test has been explained in more detail in section 3.3.3. Participants' pre-test scores were recorded automatically by Gorilla for the researcher to use later.

Four or five days later (depending on the time scheduled by each participant), the participants began the identification perception training sessions. These sessions were

[12] The full questionnaire can be found in Appendix B.

completed during a simultaneous Zoom meeting with the researcher. The three sessions were completed on three consecutive days.

Following the perception identification training sessions, participants took a perception identification post-test on Gorilla. Again, the test was completed during a simultaneous Zoom meeting with the researcher. The scores from this post-test were also recorded automatically by Gorilla.

The delayed post-test (taken six days after the post-test) was in the same format as the post-test, and scores were also recorded automatically by Gorilla.

3.5 PROCEDURES FOR DATA ANALYSIS

To achieve the objectives of the present study, a quantitative analysis was conducted. The independent variables for the present study are

1) group condition (natural stimuli for perception training, synthetic stimuli for perception training, control with no perception training)
2) testing time (T1, T2, T3)
3) the vocabulary test scores.

The dependent variables are

1) identification accuracy at the pre-test (%)
2) identification accuracy at post-test for familiar stimuli (%)
3) identification accuracy at post-test for unfamiliar stimuli (%)
4) identification accuracy at the delayed post-test for familiar stimuli (%)

Normality tests (Shapiro-Wilk) were run, and they revealed that the data set was not normally distributed (Appendix D). Due to lack of normality, all statistical tests used are non-parametric. Further information about the statistical tests used is provided in the following chapter, as the researcher addresses each research question.

4 RESULTS AND DISCUSSION

In this section, the results of the data collection will be presented and discussed. The results have been organized by research questions, beginning with RQ1: “What is the effect of natural versus synthetic stimuli for perception training with the vowel pair /æ-ɛ/?“ in Section 4.1, and then continuing on with RQ2: “How successfully are BP L2 English students able to generalize their perception training with natural versus synthetic stimuli to unfamiliar voices and carrier words?” in Section 4.2, RQ3: “How successfully are BP L2 English students able to maintain the effects of perception training with natural versus synthetic stimuli?” in Section 4.3, and finally, RQ4: “How does proficiency level impact the effectiveness of identification perception training for BP L2 English students?” in Section 4.4.

4.1 RQ1: WHAT IS THE EFFECT OF NATURAL VERSUS SYNTHETIC STIMULI FOR PERCEPTION TRAINING WITH THE VOWEL PAIR /Æ-Ɛ/?

Research has shown that it is particularly difficult for native Brazilian-Portuguese speakers to distinguish between the English vowels /æ/ and /ɛ/ (LIMA JR., 2017; RAUBER, 2006). In an effort to provide more insight regarding identification perception training for these two sounds, the researcher chose to explore the effectiveness of this type of training with natural versus synthetic stimuli. In order to do this, fifty-six participants, native Brazilian-Portuguese speakers and learners of English, were divided into three groups - those that trained with natural stimuli, those that trained with synthetic stimuli, and those that participated in a control group and did not receive identification perception training. All participants first completed a perception pre-test. Then, the two experimental groups completed three identification perception training sessions across three consecutive days. Following the training sessions, all participants completed a perception post-test.

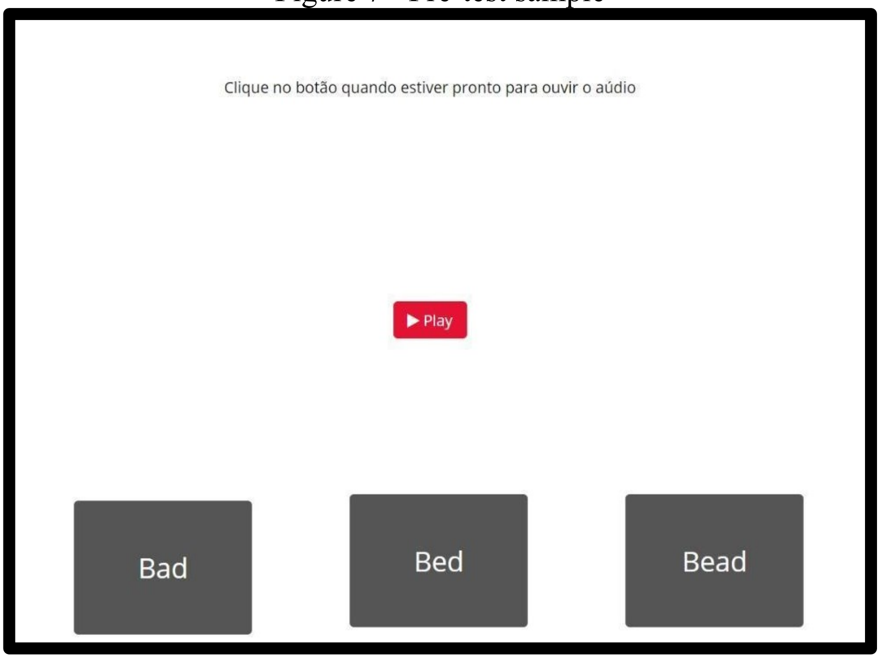
Based on the results of previous studies (NOBRE-OLIVEIRA, 2007; CHENG ET. AL, 2019; YLINEN ET.AL, 2009), the researcher hypothesized that perception training with synthetic stimuli would be the most effective. In analyzing the data, the researcher measured the participants’ pre-test and post-test scores and then compared the gain scores from pre-test to post-test for each of the three groups (natural stimuli, synthetic stimuli, control group). The results from this analysis will be shared in the following sections, beginning with the average scores from each group for the perception pre-test (Section 4.1.1.1.), the post-test scores with familiar stimuli only (Section 4.1.1.2.), the gain scores from pre-test to post-test (Section 4.1.1.3.), and the statistical significance (Section 4.1.2.) Finally, a discussion of the results will be presented in Section 4.1.4, addressing RQ1.

4.1.1 Results

4.1.1.1. Perception pre-test scores

Participants from all three groups completed the identification perception pre-test on the online platform, Gorilla Experiment Builder, during a simultaneous Zoom meeting with the researcher. They were shown triads of carrier words (each of the three words containing either the target vowel /æ/, the target vowel /ɛ/, or the distractor vowel /i/), as can be seen in Figure 7. Participants selected "play" when they were ready, and then listened to the carrier word. They were able to repeat the audio if needed. Then, they selected which of the three words on screen they heard. During the perception pre-test, participants did not receive feedback regarding their responses. Participants listened to a total of twelve different carrier words (four triads), which were each presented twice during the pre-test, for a total of twenty-four total tokens. Scores were recorded automatically by Gorilla Experiment Builder, and the researcher was able to later download those recorded results.

Figure 7 - Pre-test sample



Source - Gorilla Experiment Builder (2021)

Table 7 - Pre-test scores out of 8

	Pre-Test /æ/ Mean / 8 (SD)	Pre-Test /ɛ/ Mean / 8 (SD)
Control (n=19)	5.05 (1.58)	4.84 (1.61)
Natural (n=22)	4.82 (1.71)	4.73 (1.16)
Synthetic (n=15)	5.4 (1.24)	4.86 (2.42)

Source - The author (2021).

The average scores from the pre-test can be seen above in Table 7. As can be seen, all three groups performed higher with the /æ/ vowel than they did with the /ɛ/ vowel. The difference across vowels was bigger for the control and synthetic groups than it was for the natural group. The natural group presented fairly similar results for both vowels. The synthetic group had the highest pre-test score for both vowels.

4.1.1.2 Perception post-test scores with familiar stimuli

Following three identification perception training sessions (for the experimental groups), participants completed the identification perception post-test on Gorilla Experiment Builder during a simultaneous Zoom meeting with the researcher. The format of this test was the same as the pre-test, except that the post-test had more carrier-word triads for participants to listen to and identify (four familiar triads and three unfamiliar triads.) Below are the post-test scores for familiar stimuli only (the unfamiliar stimuli scores are not included in this table.)

Table 8 - Post-test scores out of 8 for familiar stimuli

	Post-Test /æ/ Mean / 8 (SD)	**Post-Test /ɛ/ Mean / 8 (SD)**
Control (n=19)	4.47 (1.68)	4.63 (1.07)
Natural (n=22)	6 (1.63)	5.27 (1.88)
Synthetic (n=15)	6.2 (1.74)	5.4 (1.96)

Source - The author (2021).

As can be seen in Table 8, the natural and synthetic groups performed higher for the /æ/ vowel than they did for the /ɛ/ vowel. The control group performed very similar with the two vowels, with a slightly higher average for /ɛ/. The synthetic group performed the best for both vowels out of all the groups.

4.1.1.3 Gain scores from pre-test to post-test

In order to compare the effectiveness of identification perception training with natural versus synthetic stimuli, the gain scores from the pre-test to post-test will be compared for each of the three groups (natural stimuli, synthetic stimuli, and control group.) This data can be seen below in Tables 9 and 10.

Table 9 – Averages and gain scores for familiar stimuli /æ/ out of 8

	Pre-Test /æ/ Mean / 8 (SD)	**Post-Test /æ/ Mean / 8 (SD)**	**Gain Score**
Control (n=19)	5.05 (1.58)	4.47 (1.68)	-0.58
Natural (n=22)	4.82 (1.71)	6 (1.63)	1.18
Synthetic (n=15)	5.4 (1.24)	6.2 (1.74)	0.8

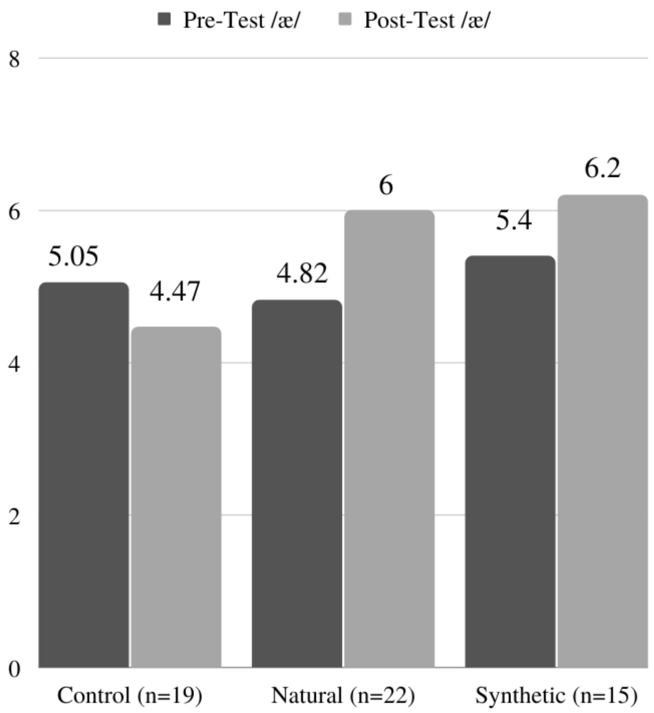
Source - The author (2021).

Table 10 – Averages and gain scores for familiar stimuli /ɛ/ out of 8

	Pre-Test /ɛ/ Mean / 8 (SD)	**Post-Test /ɛ/ Mean / 8 (SD)**	**Gain score**
Control (n=19)	4.84 (1.61)	4.63 (1.07)	-0.21
Natural (n=22)	4.73 (1.16)	5.27 (1.88)	0.54
Synthetic (n=15)	4.86 (2.42)	5.4 (1.96)	0.54

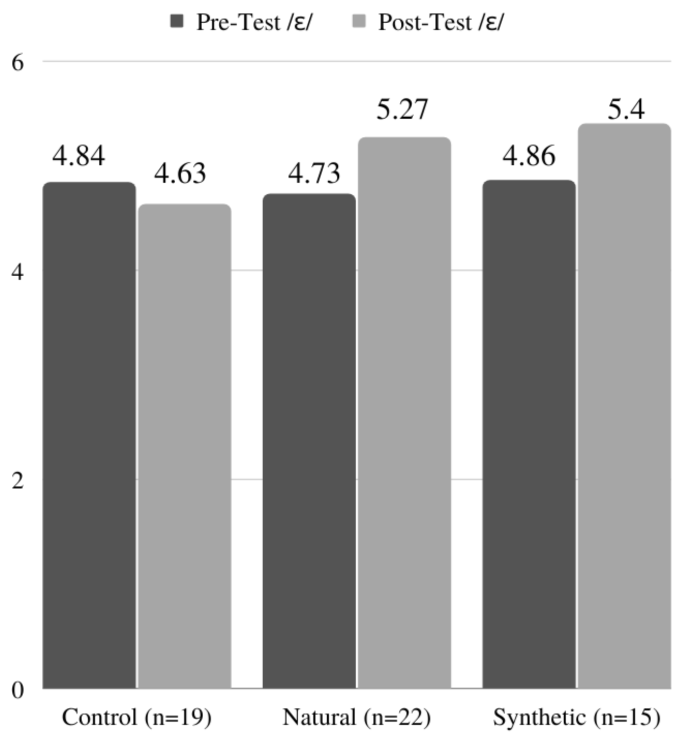
Source - The author (2021).

Figure 8 – Pre-test and post-test scores for /æ/ out of 8



Source - The author (2021)

Figure 9 – Pre-test and post-test scores for /ɛ/ out of 8



Source - The author (2021)

As can be seen in the tables and figures above, the participants in the control group performed worse in the post-test than they did in the pre-test for both vowels. However, both experimental groups performed better in the post-test compared to the pre-test. The two experimental groups also had higher gain scores for the /æ/ vowel than they did for the /ɛ/ vowel.

4.1.2 Statistical Significance

The descriptive analysis showed that the data violates assumptions for a parametric test (Appendix D), and so the non-parametric Kruskall-Wallis test was deemed appropriate for this data set and was used to compare the means for the three groups (control, synthetic and natural stimuli). As can be seen below in Table 11, the comparisons across the three groups for each type of vowel showed that only the /æ/ post-test scores were significant.

Table 11 – Kruskall Wallis test results for pre- and post-Tests

	Pre-Test /æ/	**Pre-Test /ɛ/**	**Post-Test /æ/**	**Post-Test /ɛ/**
Chi-Square	0.73	0.622	10.555	2.186
P-Value	0.694	0.733	0.005	0.335

Source - The author (2022).

As only the /æ/ post-test scores were significant, the Mann-Whitney test was then carried out for them to locate which of the three groups was performing differently. In comparing the control and natural groups in the Mann-Whitney test, the z score was -2.857, and the p-value was .004. In comparing the control and synthetic groups, the z score was -2.702, and the p-value was .007. In comparing the natural and synthetic groups, the z score was -.427, and the p-value was .669. The pairwise comparisons confirmed that the significant difference when comparing the pre and post-tests results with familiar stimuli was due to the different performances of the synthetic group (gain score = 0.8) versus the control group (-0.58), and the different performance of the natural group (1.18) versus the control group. There was no significant difference across the two control groups, showing that the treatment was equally effective for both.

A Wilcoxon test was then carried out in order to compare the pre- and post-test results for each group individually. The results from this test can be seen below in Table 12.

Table 12 – Wilcoxon test results for pre- to post-test for each group

		Pre-Test to Post-Test /æ/	Pre-Test to Post-Test /ɛ/
Control	**Z score**	-1.576	-0.343
	P value	0.115	0.732
Natural	**Z score**	-2.609	-1.765
	P value	0.009	0.078
Synthetic	**Z score**	-1.713	-0.675
	P value	0.087	0.5

Source - The author (2022).

The within-group comparisons shown in Table 12 revealed no statistical differences for the control group or the synthetic group across tests. However, it indicates that the natural group performed significantly differently with the /æ/ across tests (a gain score of 1.8, as shown in Table 9.)

4.1.3 Discussion

The results from this study indicate that identification perception training with both natural and synthetic stimuli is effective for improving native BP learners of English's ability to identify the English vowels /æ/ and /ɛ/. Participants in the control group performed worse in the post-test than they did in the pre-test, but participants in both of the experimental groups (natural and synthetic) performed better in the post-test than they did in the pre-test.

Participants in both the experimental groups had higher gain scores for the identification of the /æ/ vowel than they did for the /ɛ/ vowel. Interestingly, all groups scored higher on the pretest with the unfamiliar vowel, /æ/ than they did on the familiar vowel, /ɛ/. This may be a result of overcompensation or an effect of the testing conditions (presence of the orthography in the stimuli, for example.)

Participants in the natural stimuli group had a higher gain score for the /æ/ vowel (1.18) than the participants in the synthetic stimuli group did (0.8.) This is inconsistent with previous research (NOBRE-OLIVEIRA, 2007); however, it is important to note that the participants in the synthetic group scored higher in the pre-test than the participants in the natural group did. Therefore, it may be that the synthetic group had a lower gain score because they had less room for improvement overall. These participants also had higher total scores on the post-test than the participants in the natural group did.

4.2 RQ2: HOW SUCCESSFULLY ARE BP L2 ENGLISH STUDENTS ABLE TO GENERALIZE THEIR PERCEPTION TRAINING WITH NATURAL VERSUS SYNTHETIC STIMULI TO UNFAMILIAR VOICES AND CARRIER WORDS?

In order for perception training to be effective and applicable to real life situations, it is important for learners to be able to generalize their learning beyond just the specific voices and carrier words that they hear during the training sessions. With the goal of better understanding this aspect of perception, the researcher examined how well students who trained with synthetic versus natural stimuli were able to generalize their learning to unfamiliar voices and words. The data used for this is the same data that were reported in section 4.1; however, in order to look at generalizability, results from the unfamiliar stimuli (as well as familiar stimuli) will be examined in the post test.

In analyzing the data, the researcher compared the participants' post-test scores with familiar stimuli only to their post-test scores with unfamiliar stimuli only. These scores were then calculated as Z-scores (because there were a total of 8 familiar stimuli and 6 unfamiliar stimuli in the post-test.) Previous research has shown that learners who train with synthetic stimuli are more effective in generalizing their learning (CHENG ET. AL, 2019.) The researcher hypothesized that in this study as well, participants who trained with synthetic stimuli would be better at generalizing their learning than participants who trained with natural stimuli.

The results from this analysis will be shared in the following sections, beginning with the average scores from each group for the perception pre-test (Section 4.2.1.1), the post-test scores with the familiar stimuli only (Section 4.2.1.2.), the post-test scores with the unfamiliar stimuli only (Section 4.2.1.3.), a comparison of participants' scores from the pre-test, post-test with familiar stimuli, and post-test with unfamiliar stimuli (Section 4.2.1.3.), and the statistical significance (Section 4.2.2..) Finally, a discussion of the results will be presented in Section 4.2.4.

4.2.1 Results

4.2.1.1 Perception post-test scores with unfamiliar stimuli

The familiar stimuli and carrier words with distractor vowels have been excluded in the post-test count for unfamiliar stimuli. Results can be seen in Table 13 below. Note that the total score for the unfamiliar stimuli in the post test is 6 (different from the pre-test and the post-test with familiar items, reported in section 4.1, in which the total score for each was 8.)

Table 13 – Post-test scores (with unfamiliar stimuli only) out of 6

	Post-Test /æ/ Mean / 6 (SD)	**Post-Test /ɛ/ Mean / 6 (SD)**
Control (n=19)	3.68 (1.42)	3 (1.49)
Natural (n=22)	4.55 (1.41)	3.68 (1.21)
Synthetic (n=15)	4.53 (1.25)	3.73 (1.33)

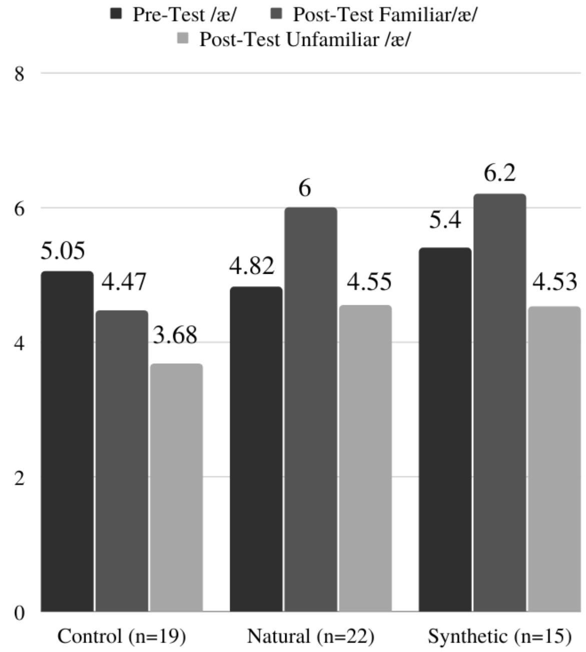
Source - The author (2021)

As can be seen above, the three groups scored higher on the post-test for the /æ/ vowel than they did for the /ɛ/ vowel. Again, this is likely due to participants' overcompensation for the difficult /æ/ sound.

4.2.1.2 Comparison of participants' scores from the pre-test, post-test with familiar stimuli, and post-test with unfamiliar stimuli scores

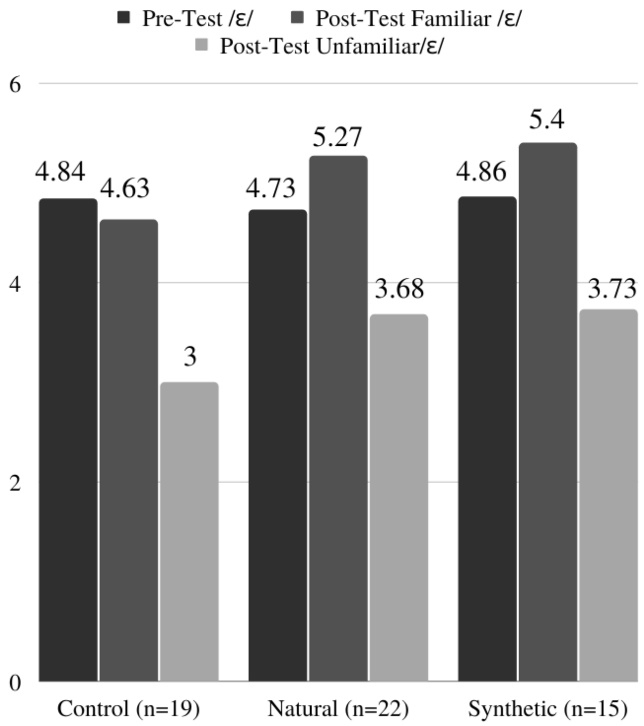
In this section, a comparison of participants' scores for familiar and unfamiliar stimuli will be shown. Below in Tables 14 and 15, participants' average scores are shown. Z scores were calculated and used to run the statistical tests for this data, however percentages are shown here, as they are easier to understand.

Figure 10 – Pre-test, post-test familiar, and post-test unfamiliar scores for /æ/



Source - The author (2021)

Figure 11 – Pre-test, post-test familiar, and post-test unfamiliar scores for /ɛ/



Source - The author (2021)

Table 14 – Pre-test, post-test familiar, and post-test unfamiliar averages for /æ/

	Pre-Test /æ/ Mean % (SD)	**Post-Test /æ/ Mean % (SD)**	**Post-Test /æ/ Mean % (SD)**
Control (n=19)	63.1% (1.58)	55.9 % (1.68)	61.3% (1.42)
Natural (n=22)	60.3% (1.71)	75% (1.63)	75.8% (1.41)
Synthetic (n=15)	67.5% (1.24)	77.5% (1.74)	75.5% (1.25)

Source - The author (2021).

Table 15 – Pre-test, post-test familiar, and post-test unfamiliar scores for /ɛ/

	Pre-Test /ɛ/ Mean % (SD)	**Post-Test /ɛ/ Mean % (SD)**	**Post-Test /ɛ/ Mean % (SD)**
Control (n=19)	60.5% (1.61)	57.9% (1.07)	50% (1.49)
Natural (n=22)	59.1% (1.16)	65.9% (1.88)	61.3% (1.21)
Synthetic (n=15)	60.8% (2.42)	67.5%	62.2% (1.33)

SOURCE - The author (2021)

As can be seen, the participants in the control group performed better with the unfamiliar stimuli than they did with the familiar stimuli for the /æ/ vowel, but performed worse with the unfamiliar stimuli than they did with the familiar stimuli for the /ɛ/ vowel. Participants in the natural group performed slightly better with the unfamiliar stimuli than they did with the familiar stimuli for the /æ/ vowel, and worse with the unfamiliar stimuli than they did for the familiar stimuli for the /ɛ/ vowel. Participants in the synthetic group performed worse with the unfamiliar stimuli than they did for the familiar stimuli for both vowels.

Participants in the control group performed worse for both familiar and unfamiliar stimuli in the post-test than they did in the pre-test for both vowels. Participants in both experimental groups (natural and synthetic) performed better for both familiar and unfamiliar stimuli in the post-test than they did in the pre-test for both vowels.

4.2.2 Statistical Significance

The descriptive analysis showed that the data violates assumptions for a parametric test (Appendix D), and so the non-parametric Kruskall-Wallis test was deemed appropriate for this data set.

Table 16 – Kruskall-Wallis results for familiar and unfamiliar stimuli in the post-test

	Familiar /æ/	Familiar /ɛ/	Unfamiliar /æ/	Unfamiliar /ɛ/
Chi-Square	10.555	2.186	4.818	2.835
P-Value	0.005	0.335	0.09	0.242

Source - The author (2022).

Wilcoxon tests were then carried out in order to compare the post-test familiar stimuli scores with the post-test unfamiliar stimuli scores for each group individually. The results from this test can be seen in Table 16.

Table 17 – Wilcoxon results for familiar versus unfamiliar stimuli in the post test for each group

		Familiar vs Unfamiliar /æ/	Familiar vs Unfamiliar /ɛ/
Control	**Z**	-1.884	-3.449
	P value	0.06	0.001
Natural	**Z**	-2.792	-3.100
	P value	0.005	0.002
Synthetic	**Z**	-2.394	-2.386
	P value	0.017	0.017

Source: The author (2022).

These results demonstrate that all groups performed significantly differently when identifying familiar and unfamiliar vowel tokens.

4.2.3 Discussion

The results indicate that identification perception training with both natural and synthetic stimuli is effective for generalizable perception learning. Although participants in the control group had a lower percentage score for the unfamiliar stimuli in the post test than they did in the pre-test, participants from both of the experimental groups (synthetic and natural) had a higher percentage score for the unfamiliar stimuli in the post test than they did in the pre-test. The natural group seems to have done a better job generalizing their learning for both vowels than the synthetic group did.

Participants from all groups had higher pre-test and unfamiliar post test scores for the /æ/ vowel than they did for the /ɛ/ vowel. As mentioned previously, this may be due to overcompensation by the participants for the difficult vowel sound. However, both of the experimental groups performed higher for the /ɛ/ vowel in the post-test with familiar stimuli.

4.3 RQ3: HOW SUCCESSFULLY ARE BP L2 ENGLISH STUDENTS ABLE TO MAINTAIN THE EFFECTS OF PERCEPTION TRAINING WITH NATURAL VERSUS SYNTHETIC STIMULI?

Aside from the immediate effectiveness and generalizability of perception training, it is also important to consider whether students are able to retain their learning following perception training. The researcher sought to understand whether the results of identification perception training with synthetic and natural stimuli were maintained after a 6 day period. To do so, the fifty-six participants completed a delayed post-test following the pre-test, three perception training sessions (experimental groups only), and the post test. The delayed post-test was in the same format and used the same stimuli as the post-test.

In analyzing the data, the researcher measured the pre-test scores, post-test scores (familiar stimuli only), and delayed post-test scores (familiar stimuli only.) The results from this analysis will be shared in the following sections, beginning with the delayed post-test scores with the familiar stimuli only (Section 4.3.1.1.), a comparison of the pre-test, post-test and delayed post-test scores (Section 4.3.2.), and the results from the statistical inference tests (Section 4.3.3.). Finally, a discussion of the results will be presented in Section 4.3.4.

4.3.1 Results

4.3.1.1 Delayed post-test scores

Six days after taking the post-test, participants from all three groups (natural, synthetic, and control) completed the delayed perception post-test. This test was identical to the post-test, with the same stimuli and format. The only difference was the order in which the triads of carrier words were presented to the participants, which was automatically randomized by the platform, Gorilla Experiment Builder. This platform automatically recorded participants' results, which the researcher later downloaded and scored. The results can be seen in Table 18.

Table 18 – Delayed post test scores (familiar stimuli)

	Delayed Post-Test /æ/ Mean / 8 (SD)	**Delayed Post-Test /ɛ/ Mean / 8 (SD)**
Control (n=19)	4.95 (1.35)	5.68 (1.67)
Natural (n=22)	5.36 (2.06)	5.14 (1.96)
Synthetic (n=15)	5.87 (2.10)	5.80 (1.93)

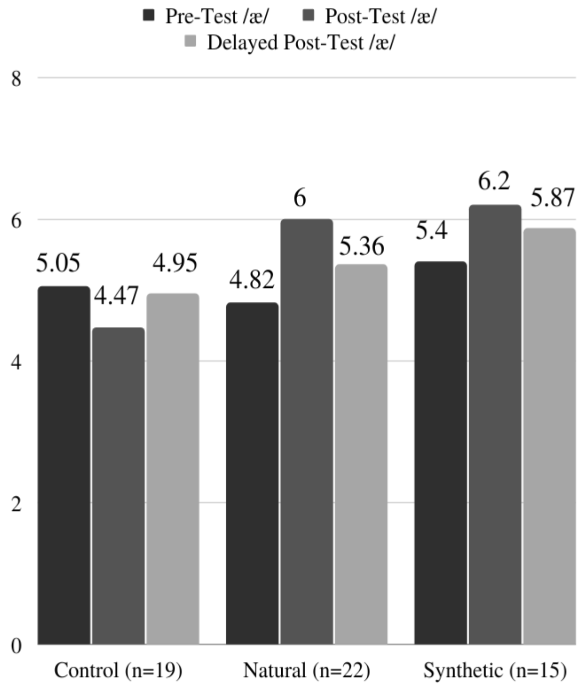
Source - The author (2021).

As can be seen above, the participants in the control group scored higher for the /ɛ/ vowel than they did for the /æ/ vowel. Participants in the natural and synthetic groups both scored higher for the /æ/ vowel than they did for the /ɛ/ vowel. This may be due to overcompensation by the participants for a difficult sound. Of the three groups, the synthetic group performed the best for both vowels overall.

4.3.1.2 Comparing the Pre-Test, Post-Test, and Delayed Post-Test Scores

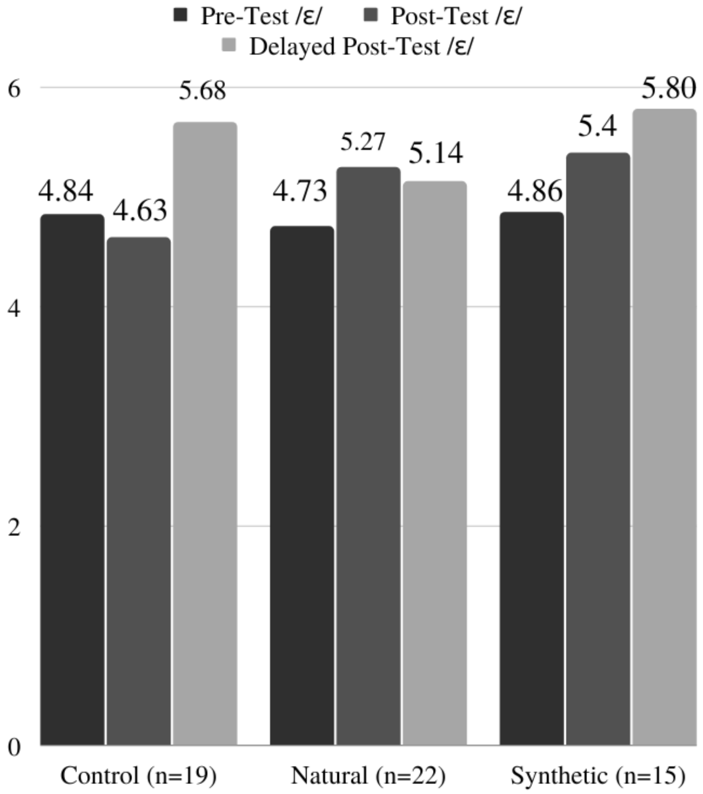
The average scores for the pre-test, post-test, and delayed post-test can be seen for each of the three groups in Figures 12 and 13 below.

Figure 12 – Pre-test, post-test, and delayed post-test scores for /æ/



Source - The author (2021)

Figure 13 – Pre-test, post-test, and delayed post-test scores for /ɛ/



Source - The author (2021)

Table 19 – Pre-test, post-test, and delayed post-test /ɛ/

	Pre-Test /ɛ/ Mean / 8 (SD)	**Post-Test /ɛ/ Mean / 8 (SD)**	**Delayed Post-Test /ɛ/ Mean / 8 (SD)**
Control (n=19)	4.84 (1.61)	4.63 (1.07)	5.68 (1.67)
Natural (n=22)	4.73 (1.16)	5.27 (1.88)	5.14 (1.96)
Synthetic (n=15)	4.86 (2.42)	5.4 (1.96)	5.80 (1.93)

Source - The author (2021).

Table 20 – Pre-test, post-test, and delayed post-test /æ/

	Pre-Test /æ/ Mean / 8 (SD)	**Post-Test /æ/ Mean / 8 (SD)**	**Delayed Post-Test /æ/ Mean / 8 (SD)**
Control	5.05 (1.58)	4.47 (1.68)	4.95 (1.35)

	Pre-Test /æ/ Mean / 8 (SD)	Post-Test /æ/ Mean / 8 (SD)	Delayed Post-Test /æ/ Mean / 8 (SD)
(n=19)			
Natural (n=22)	4.82 (1.71)	6 (1.63)	5.36 (2.06)
Synthetic (n=15)	5.4 (1.24)	6.2 (1.74)	5.87 (2.10)

Source - The author (2021).

In comparing the results from the pre-test (Section 4.1.1.1), post-test (Section 4.1.1.2) and delayed post-test, it is interesting to note that the participants in the control group performed worse on the post-test than they did on the pre-test for both vowels, /ɛ/ and /æ/ (Tables 19 and 20). However, these participants then improved on the delayed post-test. For the /ɛ/ vowel, these participants performed even better on the delayed post-test than they did on the pre-test. As for the two experimental groups, participants from both groups performed better on the post-test than they did on the pre-test for both vowels. Participants in the natural group were able to maintain some of their learning after a 6-day period for both vowels, as their delayed post-test scores were better than their pre-test scores (but worse than their post-test scores). Participants in the synthetic group were also able to maintain their learning. For the /ɛ/ vowel, participants in the synthetic group actually performed better on the delayed post-test than they did on either the pre-test or the post-test. For the /æ/ vowel, they performed similarly to the natural group. They maintained some of their learning, as their delayed post-test scores were higher than their pre-test scores. However, like the natural group, they performed worse on the delayed post-test than they did on the post-test.

4.3.2 Statistical significance

The descriptive analysis showed that the data violate assumptions for a parametric test (Appendix D), and so the non-parametric Kruskall-Wallis test was deemed appropriate for this data set. The only comparison that was significant here was the /æ/ post-test familiar stimuli scores, and the Mann-Whitney test was already carried out on this data set, as can be seen in Section 4.1.2. Thus, the results indicate that the differences in mean observed for the delayed posttest are not significant for any of the three groups.

Table 21 – Kruskall-Wallis results for the post-test and the delayed post-test

	Post-Test /æ/	Post-Test /ɛ/	Delayed Post-Test /æ/	Delayed Post-Test /ɛ/
Chi-Square	10.555	2.186	1.486	3.389
P-Value	0.005	0.335	0.476	0.184

Source- The author (2022).

A Wilcoxon test was then carried out in order to compare the post-test with the delayed post-test scores for each group individually. The results from this test can be seen in Table 22.

Table 22 – Wilcoxon results for the post-test versus the delayed post-test for each group

		Post-Test to Delayed Post-Test /æ/	Post-Test to Delayed Post-Test /ɛ/
Control	**Z score**	-0.826	-2.256
	P value	0.409	0.024
Natural	**Z score**	-1.341	-.201
	P value	0.18	0.841
Synthetic	**Z score**	-0.763	-0.714
	P value	0.446	0.475

Source- The author (2022).

These results show that the differences in mean observed across the post-test and the delayed post-test were significant for the control group when identifying the /ɛ/ vowel only.

4.3.3 Discussion

The results indicate that identification perception training is effective not only for learning, but also for retaining learning. Participants maintained some of their learning for both vowels, as most of their delayed post-test scores were better than their pre-test scores (but worse than their post-test scores.) Notably, participants in the synthetic group actually performed better on the delayed post-test than they did on either the pre-test or the post-test for the /ɛ/ vowel.

The control group's average score decreased from pre-test to post-test; however, it actually increased in the delayed post-test. This may be a practice effect, as by the delayed post-test, the participants had received more exposure to the format of the test and the stimuli. These participants also had less exhaustion from the repetitive online tasks than the participants in the experimental groups did. It should be pointed out that the inferential tests showed no significant effect for most of the within-group comparisons, except for the post-test and the delayed post-test for the control group when identifying the /ɛ/ vowel.

4.4 RQ4: HOW DOES PROFICIENCY LEVEL IMPACT THE EFFECTIVENESS OF IDENTIFICATION PERCEPTION TRAINING FOR BP L2 ENGLISH STUDENTS?

4.4.1 Results

In addition to exploring the use of perception training with natural versus synthetic stimuli, the generalizability of learning, and the delayed results, the researcher was interested in whether students' L2 proficiency level affected the effectiveness of perception training. The researcher hypothesized that students' with a higher level of proficiency would benefit more from the perception training, as these students would be less likely to deal with spelling or low-frequency word challenges as lower level students might. In order to test this hypothesis, the researcher looked for a relationship between students' L2 proficiency level and their gain scores from pre-test to post-test.

The results from this analysis will be shared in the following sections, beginning with the proficiency test scores (Section 4.4.1.1.), the gain scores from pre-test to post-test (Section 4.4.1.2.), and the results from the Mann Whitney test (Section 4.4.2.) Finally, a discussion of the results will be presented in Section 4.4.4.

4.4.1.1 Proficiency level

In order to measure participants' proficiency level, the Lognostics online Yes/No vocabulary test was carried out. In this test, participants were presented with a series of 200 vocabulary words, one at a time. The maximum possible score was 10,000. Based on the manual's description of the scores, the researcher labeled the participants proficiency level in the following manner:

Score below 2,500Not included in analysis

Score of 2,500-3,500Beginner

Score of 3,501 - 6,000Intermediate

Score of 6,001 - 10,000Advanced

Following the recommendation of the Lognostics Yes/No Vocabulary Test manual, the data for Participants P117 and P138 have not been included in this analysis, as their scores were below 2,500. The manual noted that, "scores below 2,500 words are probably unreliable and should be treated with extreme caution" (MEARA, MIRALPEIX, 2014.)

4.4.2 Statistical significance

The descriptive analysis showed that the data for the vocabulary test do not violate assumptions for the parametric test (Appendix D), However, as the data for the perception tests was found to be not normally distributed, nonparametric correlation (Spearman) tests were performed. The results from the Spearman test can be seen below in Table 23. Based on these test results, it appears that proficiency level had a significant correlation to performance in some of the perception tests for the control group. As all the significant correlations were positive, it seems that the more proficient learners in the control group performed better than the less proficient learners in the perception identification test with /ɛ/ in the pre-test and the delayed post-test, and with /æ/ in the post-test. However, the only significant correlation for the experimental groups was between proficiency level and perception of /ɛ/ in the pre-test for the synthetic group only, thus indicating that in this group, the more proficient learners performed better with the /ɛ/ vowel than the less proficient learners in the pre-test.

Table 23 – Spearman correlation coefficients

	Vocabulary score - control	**Vocabulary score - natural**	**Vocabulary score - synthetic**
Pre-test /æ/	.410	.280	.131
Pre-test /ɛ/	.459*	.109	.567*
Post-test /æ/	.541*	.130	.121
Post-test /ɛ/	-.205	.187	.495
Delayed post-test /æ/	-.350	.299	.215
Delayed post-test /ɛ/	.539*	.407	.300

* p= < 0.05

Source - The author (2022).

4.4.3 Discussion

From the data, it is difficult to come to a conclusion regarding the impact that proficiency level has on the effectiveness of identification perception training. There were only significant correlations between proficiency level and performance in the perception tests for the control group and for the pre-test /ɛ/ and proficiency level for the synthetic group.

The vocabulary test used may not be the most accurate indicator of participants' true L2 English proficiency. As the test was online and repetitive, some participants may have been affected by exhaustion. Also, the test rated proficiency based on vocabulary knowledge, whereas a perception test may have been more appropriate for this project.

It's also important to consider the individual differences among participants, which could have also impacted results. Some participants reported having studied other languages, spending extended periods of time abroad, and varying methods and environments for learning English prior to participating in the data collection process, which may have also contributed to their successfulness or lack thereof with the perception training.

5 FINAL REMARKS

In this final chapter, the results from this study will be summarized (Section 5.1.) The limitations of the present study will be noted in Section 5.2 and some suggestions for further research will be described in Section 5.3. Finally, the chapter will conclude with a discussion of the implications the results from this study provide in Section 5.4.

5.1 CONCLUSION

The findings of this study will be summarized in the following paragraphs, organized by research question.

5.1.1 RQ1: What is the effect of natural versus synthetic stimuli for perception training with the vowel pair /æ-ɛ/?

The results from the present study indicate that identification perception training with both natural and synthetic stimuli is effective for improving native BP learners of English's ability to identify the English vowels /æ/ and /ɛ/. Participants in the control group performed worse in the post-test than they did in the pre-test, but participants in both experimental groups (natural and synthetic) performed better in the post-test than they did in the pre-test after participating in three consecutive identification perception training sessions. The results from this study suggest that perhaps the use of natural stimuli is more effective for Brazilian learners' improvement in the identification of the English vowels /æ-ɛ/, although the data were not statistically significant for most comparisons, and this is not consistent with some previous research (NOBRE-OLIVEIRA, 2007.)

5.1.2 RQ2: How successfully are BP L2 English students able to generalize their perception training with natural versus synthetic stimuli to unfamiliar voices and carrier words?

The results from this study indicate that identification perception training with both natural and synthetic stimuli is effective for generalizable perception learning. Although participants in the control group had a lower percentage score for the unfamiliar stimuli in the post test than they did in the pre-test, participants from both experimental groups (synthetic and natural) had a higher percentage score for the unfamiliar stimuli in the post test than they did in the pre-test. The results from this study suggest that perhaps the use of natural stimuli is more effective for learners' ability to generalize their learning, although the data were not statistically significant.

5.1.3 RQ3: How successfully are BP L2 English students able to maintain the effects of perception training with natural versus synthetic stimuli?

The results from this study indicate that identification perception training is effective not only for learning, but also for retaining learning. Participants maintained some of their learning for both vowels, as most of their delayed post-test scores were better than their pre-test scores (but worse than their post-test scores.) Notably, participants in the synthetic group actually performed better on the delayed post-test than they did on either the pre-test or the post-test for the /ɛ/ vowel.

5.1.4 RQ4: How does proficiency level impact the effectiveness of identification perception training for BP L2 English students?

From the data, it is difficult to come to a conclusion regarding the impact that proficiency level has on the effectiveness of identification perception training. There was a tendency for proficiency to play a positive role in the performance of both the control group and the experimental groups, given that most correlations were positive. However, no significant correlations were obtained for the experimental groups in the post-tests, which indicates that individual differences need to be considered in studies involving L2 perception training. As mentioned in Section 4.4, the individual differences of participants and the vocabulary test used to evaluate proficiency may have affected results.

5.2 LIMITATIONS OF THE STUDY

One of the primary limitations of the study was the training and testing conditions for the participants. As a result of the Covid-19 pandemic, data collection was necessarily done online. This made it difficult to control for things such as the type of earphones used by participants or the surrounding noise and other distractions in their physical environment.

Aside from the inability to control the environment where testing and training took place, the participants also likely experienced exhaustion from the repetitive activity, which took place over the course of several consecutive days. This may explain why the control group had a higher gain score from post-test to delayed post-test than the experimental groups did, seeing as the control group had less exhaustion (as they did not participate in as many sessions.)

The participants for this study were recruited online via social media. Although this allowed for a larger sample size for the study, it also made controlling for participants' L2 English studies difficult. Some participants were actively enrolled in L2 language programs at the time of data collection, whereas others were not. Those that were actively studying at the time of data collection reported a variety of methods and environments. Some participants reported studying with a private teacher, others enrollment in an L2 language school, while others reported self-study.

Ideally, the researcher would have liked to provide more training sessions, extended the delayed post-test to several months after the post-test (instead of just one week after,) utilized a larger variety of carrier words, and utilized a larger variety of voices and accents for the stimuli creation. However, the decisions to carry out the study in this way were made as a result of the limited time for a master's project.

5.3 SUGGESTIONS FOR FURTHER RESEARCH

Continued research to compare the use of natural versus synthetic stimuli for perception training is needed, in order to better understand whether one of these two is more effective. Studies over a longer period of time, with more training sessions and a more controlled training environment may be successful in evaluating this.

Based on the results of this study, another suggestion for further research is to more extensively investigate the effectiveness of proficient L2 English speakers' voices for perception training (rather than native-speakers' voices only.) Considering the now wide use of English as a lingua-franca, this adaptation to what has been standard for perception training seems relevant. The researcher suggests the usage of speakers from various geographical regions with various accents.

5.4 IMPLICATIONS

The results of this study confirm the effectiveness of identification perception training (with both natural and synthetic stimuli) for improving Brazilian learners' identification of the English vowel sounds /æ-ɛ/. This study also suggests that the use of native Brazilian-Portuguese L2 English speakers' voices for stimuli can be effective for perception training, and that learning from training with these voices is generalizable to new voices and words. Another implication from this study is that controlled tasks alone can trigger attention and lead to learning. The correlational results from this study also suggest that perception training is not exclusively beneficial for learners with an already high L2 proficiency level. The results support SLM (FLEGE, 1995) as well, as they imply that new phonetic categories can be formed through training.

REFERENCES

ALIAGA-GARCÍA, Cristina; MORA, Joan C. Assessing the effects of phonetic training on L2 sound perception and production. *In:* **Recent research in second language phonetics/phonology: perception and production.** Newcastle upon tyne: Cambridge Scholars Publishing, 2009. p. 2-32.

ANWYL-IRVINE, A.L; MASSONIÉ, J.; FLITTON, A.; KIRKHAM, N.Z.; EVERSHED, J.K. **Gorilla experiment builder**, 2018. Available at: https://gorilla.sc/. Access on: 01 jun 2021.

ALVES, Ubiratã Kickhöfel.; LUCHINI, Pedro Luis. Effects of perceptual training on the identification and production of word-initial voiceless stops by Argentinean learners of English. **Ilha do Desterro,** [S.L.], v. 70, n. 3, (pp. 15-32). Universidade Federal de Santa Catarina (UFSC), 2017. https://doi.org/10.5007/2175-8026.2017v70n3p15.

BANIK, Shipro; KIBRIA, BM; SHARMA, D. Testing the population coefficient of variation. **Journal of modern applied statistical methods**, [*S. l.*], v. 11, n. 2, p. 325-335, nov. 2012. DOI 10.22237/jmasm/1351742640. Available at: https://www.researchgate.net/publication/288656530_Testing_the_Population_Coefficient_of_Variation. Access on: 6 jun. 2022.

BARRIUSO, Taylor Anne; HAYES-HARB, Rachel. (2018) High Variability Phonetic Training as a bridge from research to practice. **The CATESOL Journal.** v. 30, n. 1, p. 177-194.

BEST, Catherine; TYLER, Michael Douglas. Nonnative and second-language speech perception: Commonalities and complementarities. *In*: MUNRO, Murray; BOHN, Ocke-Schwen (ed.). **Language experience in second language speech learning**: In honor of James Emil Flege. Amsterdam, Netherlands: 2007. cap. 2, p.13-14.

CARLET, Angelica; CEBRIAN, Juli. Identification vs. discrimination training: Learning effects for trained and untrained sounds. *In:* **Proceedings of the 18th International Congress of Phonetic Sciences.** 2015, Glasgow, Scotland. Available at: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0752.pdf.

CARLET, Angelica; CEBRIAN, Juli. Training Catalan speakers to identify L2 consonants and vowels: A short-term high variability training study. *In:* **Proceedings of the 7th International Symposium on the Acquisition of Second Language Speech.** Concordia University Working Papers in Applied Linguistics., 2014. p. 85-98.

COSTA, Bruno Coriolano de Almeida. **Pronunciation teaching is not a one-size-fits-all endeavor:** EFL teachers' beliefs and classroom practices. Florianópolis: PPGI/UFSC, 2016 (MA thesis).

CEBRIAN, Juli; CARLET, Angelica.; GORBA, Celia.; GAVALADA, Nuria. (2019). Perceptual training affects L2 perception but not cross-linguistic similarity. *In:* **Proceedings of the 19th International Congress of Phonetic Sciences**, 2019, Melbourne, Australia. Available at: https://assta.org/proceedings/ICPhS2019/papers/ICPhS_978.pdf.

CHENG, Bing; ZHANG, Xiaojuan; FAN, Siying.; ZHANG, Yang. (2019) The role of temporal acoustic exaggeration in high variability phonetic training: A behavioral and ERP study. **Frontiers in Psychology**, v. 10, 24 may 2019. https://doi.org/10.3389/fpsyg.2019.01178.

COLANTONI, Laura.; STEELE, Jeffrey.; ESCUDERO, Paola. **Second language speech:** Theory and practice. Cambridge: Cambridge University Press, 2015.

ESCUDERO, P.; BOERSMA P.; SCHURT RABUER, A.; BION, R.A.H. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *In:* **The Journal of the Acoustical Society of America***,* 2009, v. 126, n. 3, p. 1379-1393.

FLEGE, James Emil. Assessing constraints on second-language segmental production and perception. *In*: A., Meyer; N., Schiller (ed.). **Phonetics and phonology in language comprehension and production:** Differences and similarities. Berlin: Mouton de Gruyter, 2003. p. 319-355. Available at: http://jimflege.com/chapters_after_2000.html. Access on: 21 oct. 2019

FLEGE, James Emil. Two procedures for training a novel second language phonetic contrast. **Applied Psycholinguistics**, The United States of America, v**.** 16, p. 425–442. 1995. Available at: https://www.semanticscholar.org/paper/Two-procedures-for-training-a-novel-second-language-Flege/39dcc2e43c3dc10cb71f17352afedb539c2d1c57.

FLEGE, James Emil; BOHN, Ocke-Schwen. The Revised Speech Learning Model (SLM-r). *In:* RATREE, Wayland. **Second Language Speech Learning**: Theoretical and Empirical Progress, 2021.

HILLENBRAND, James; GETTY, Laura A..; CLARK, Michael J..; WHEELER, Kimberlee. Acoustic characteristics of American English vowels. **The Journal of the Acoustical Society of America**, v. 97, p. 3099-3111. 1995. https://doi.org/ 10.1121/1.4 11872.

IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.

KUHL, Patricia K..; CONBOY, Barbara T..; COFFEY-CORINA, Sharon.; PADDEN, Denise; RIVERA- GAXIOLA, Maritza.; NELSON, Tobey. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). **Philosophical Transactions of the Royal Society**, v. 363, p. 979-1000. 2007. DOI: 10.1098/rstb.2007.2154

LADEFOGED, Peter. **Vowels and consonants.** 2. ed. Oxford, UK: Blackwell Publishing, 2007.

LIMA JUNIOR, Ronaldo Mangueira. The influence of metalinguistic knowledge of segmental phonology on the production of English vowels by Brazilian undergraduate students. **Ilha do Desterro**, v. 70, n. 3, p. 117-130. 2017, Universidade Federal de Santa Catarina. https://doi.org/10.5007/2175-8026.2017 v70n3p117.

LIVELY, Scott E.; PISONI, David B.; YAMADA, Reiko A.; TOHKURA, You'ichi; YAMADA, Tsuneo. Training Japanese listeners to identify English /r/ and /l/ III: Long-term retention of new phonetic categories. *In* **The Journal of the Acoustical Society of America**. v 96, n. 4, p. 2076-2087, 1994.

LIVELY, Scott E.; LOGAN, John, S.; PISONI, David, B. (1993) Training Japanese listeners to identify English /r/ and /l/ II: The role of phonetic environment and talker variability in learning new perceptual categories. *In* **The Journal of the Acoustical Society of America**. v 94, n. 3, p. 1242-1255.

LOGAN, John S.; LIVELY, Scott E.; PISONI, David B. (1991) Training Japanese listeners to identify English /r/ and /l/: A first report. *In* **The Journal of the Acoustical Society of America**. v 80, n. 2, p. 874-886.

MACDONALD, Shem. Pronunciation – views and practices of reluctant teachers. **Prospect**, v. 17, p. 3-18. 2002, La Trobe University.

MEARA, Paul M.; JONES, Glyn. **The Eurocentres Vocabulary Size Test:** Version E1. 1, K10, MSDOS. Zurich: Eurocentres. 1989.

MEARA, Paul; MIRALPEIX, Imma. **Lognostics**, 2014. Available at: https://www.lognostics.co.uk/tools/. Access on: 01 jun 2021.

NOBRE-OLIVEIRA, Denize. **The effect of perceptual training on the learning of English vowels by Brazilian Portuguese speakers.** Florianópolis: PPGI/UFSC, 2007 (Doctoral dissertation).

NOZAWA, Tasheki. Effects of training methods and attention on the identification and discrimination of American English coda nasals by native Japanese listeners. *In* **Proceedings of the 18th International Congress of Phonetic Sciences.** Glasgow, UK: The University of Glasgow, 2015.

POLLACK, Irwin; PISONI, David. On the comparison between identification and discrimination tests in speech perception. **Psychonomic Science**, Ann Arbor: The University of Michigan, v. 24, n. 6, p. 299-300, 24 october 2013. DOI: https://doi.org/10.3758/BF03329012. Available at: https://link.springer.com/article/10.3758/BF03329012. Access on: 6 jun 2022.

QIAN, Manman; CHUKHAREV-HUDILAINEN, Evgeny; LEVIS, John. A system for adaptive high-variability segmental perceptual training: Implementation, effectiveness, transfer. **Language, Learning, and Technology**, v. 22, n. 1, p. 69-96, 2018.

RATO, Anabela. Effects of perceptual training on the identification of English vowels by native speakers of European Portuguese. **Concordia Working Papers in Applied Linguistics,** v. *5*. p. 529-546, 2014.

RAUBER, Andreia. **Perception and production of English vowels by Brazilian EFL speakers.** Florianópolis: UFSC, 2006 (Doctoral dissertation).

SEARA, Izabel Christine; NUNES, Vanessa Gonzaga; LAZZAROTTO-VOLCÃO, Cristiane. **Fonética e fonologia do português brasileiro.** São Paulo, Brasil: Editora Contexto, 2015.

SHINOHARA,Yasuaki; IVERSON, Paul. High variability identification and discrimination training for Japanese speakers learning English /r/–/l/. **Journal of Phonetics** v. 66, p. 242-251, 2018. https://doi.org/10.1016/j.wocn.2017.11.002.

SILVA, Thais Cristofaro; SEARA, Izabel; SILVA, Adelaide.; RAUBER, Andreia Schurt; CANTONI, Maria. **Fonética Acústica: Os sons do português brasileiro**. São Paulo, Brasil: Editora Contexto, 2019.

STOEHR, Antje; MARTIN, Clara D. Orthography affects L1 and L2 speech perception but not production in early bilinguals. *In:* Jubin, A., Clahsen, H. (ed.) **Bilingualism: Language and Cognition**, v 25, p. 108-120, 2021. https://doi.org/10.1017/ S1366728921000523

STRANGE, Winifred.; DITTMAN, Sibylla. Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. **Perception and Psychophysics**, v. 36, n. 2, p. 131-145, 1984.

WANG, Xinchun; MUNRO, Murray J. The perception of English tense-lax vowel pairs by native Mandarin speakers: The effect of training on attention to temporal and spectral cues. **Proceedings of the 14th International Congress of Phonetic Sciences**, p. 125-128. San Francisco, 1999.

YLINEN, Sari; UTHER, Maria; LATVALA, Antii.; VEPSALAINEN, Sara.; IVERSON, Paul.; AKAHANE-YAMADA, Reiko; NAATANEN, Risto. Training the brain to weight speech cues differently: A study of Finnish second-language users of English. **Journal of cognitive neuroscience**. v. 22, p. 1319-1332, 2009.

APPENDIX A - CONSENT FORM

Universidade Federal de Santa Catarina
Centro de Comunicação e Expressão
Programa de Pós-Graduação em Inglês: Estudos Linguísticos e Literários

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado(a) aluno(a),

Meu nome é **Elisabeth Ann Bunch Oliveira da Rosa** e sou mestranda no Programa de Pós-Gradução em Inglês (PPGI) da Universidade Federal de Santa Catarina (UFSC). Junto com a minha orientadora , Prof. Dra. Rosane Silveira, estamos o(a) convidando a participar de uma pesquisa científica intitulada: *The effect of perception training with synthetic and natural stimuli on BP learners' ability to discriminate the English vowels /æ-ɛ/.*

1) OBJETIVOS DA PESQUISA:

O principal objetivo dessa pesquisa é investigar a efetividade de estímulos naturais e sintéticos para o
treinamento de percepção com as vogais /æ-ɛ/ em inglês para estudantes brasileiros de inglês, dos níveis iniciante e avançado. Também, essa pesquisa busca entender a capacidade de alunos a generalizar a aprendizagem no treinamento de percepção a novos estímulos e entender os efeitos a longo prazo de treinamento de percepção.

2) PROCEDIMENTOS DA COLETA DE DADOS:

A coleta de dados se fará através de testes de percepção com TP 3.1. software e com um questionário
escrito. É importante relatar que durante toda a coleta de dados, a pesquisadora estará disponível para ajudar você com possíveis problemas ou dificuldades que ocorrerem.

Ao primeiro encontro, você responderá ao um questionário escrito sobre sua experiência na aprendizagem de inglês e outras línguas. Em seguida, você realizará um teste de percepção com TP 3.1. software.

Depois da realização deste teste, você participará de quatro sessões de treinamento de percepção de vogais em inglês. Estas sessões durarão aproximadamente dez minutos cada, e serão realizadas durante o seu horário de aula. Serão feitos durante quatro aulas seguidas.

Ao completar essas quatro sessões de treinamento de percepção, você realizará um outro teste de percepção no sexto encontro com a pesquisadora.

Aproximadamente um mês depois do sexto encontro, você realizará mais um teste de percepção para completar o sétimo e último encontro com a pesquisadora.

3) POSSÍVEIS BENEFÍCIOS, GANHOS E RESULTADOS DA PESQUISA:

Você terá benefício de medir sua percepção das vogais /æ-ɛ/ em inglês. Pode também possivelmente melhorar sua percepção destes e outras vogais em inglês. Além disso, um relatório final com os resultados da pesquisa e dos testes serão enviados para você por e-mail assim que forem concluídos.

4) POSSÍVEIS RISCOS E DESCONFORTOS DA PESQUISA:

Toda interação humana pode apresentar riscos e desconfortos para alguém. Assim, essa pesquisa também
pode apresentar alguns riscos e desconfortos, como: aborrecimento, fadiga, constrangimento, cansaço, desconforto físico ou psicológico, ansiedade, e mudança na autoestima.

5) CONFIDENCIALIDADE DA IDENTIDADE E INFORMAÇÕES:

Este documento garante a confidencialidade da sua identidade e informações privadas. Ou seja, a garantia
de que as informações privadas estão protegidas e confiadas as pesquisadoras que tomarão todas as providências necessárias para manter o sigilo. As informações não serão reveladas sem as devidas autorizações. Porém, sempre existe a possibilidade da quebra de sigilo, mesmo que não intencional e/ou involuntária, cujas consequências serão tratadas nos termos da lei.

6) POSSÍVEL IDENIZAÇÃO:

Caso haja algum dano material ou imaterial devidamente comprovado da pesquisa, este documento
garante o reparo ao dano que deve ser pago de acordo com a Resolução 510/16.

7) DESPESAS NA PESQUISA:

Como a pesquisa será realizada no seu horário de aula, não se faz necessário nenhum tipo de
ressarcimento. Todo material utilizado será custeado pelas pesquisadoras.

8) DESISTÊNCIA DA PARTICIPAÇÃO NA PESQUISA:

Caso você não queira continuar a sua participação na pesquisa ou que os dados coletados não sejam
usados, não há nenhum problema. A desistência pode ocorrer a qualquer momento, sem qualquer prejuízo. Basta entrar em contato comigo através do número de whatsapp ou por e-mail.

9) ASSISTÊNCIA, CONTATOS E ENDEREÇO DOS PESQUISADORES:

Ao longo da pesquisa, você receberá acompanhamento e assistência necessários caso haja alguma dúvida ou problema. Abaixo os contatos e endereço das pesquisadoras para tirar qualquer dúvida ou pedir mais informações:

Pesquisadora: **Elisabeth Ann Bunch Oliveira da Rosa**
Whatsapp: **(+1) 360 224 9148**
E-mail: **elisabeth.bunch@ymail.com** , **elisabethbunch@gmail.com**

Endereço do Programa de Pós-Graduação em Inglês (PPGI):
Centro de Comunicação e Expressão – CCE "B" Sala 313
Campus Universitário – Trindade – Florianópolis – SC
CEP: 88.040-900
Orientadora e Doutora: **Rosane Silveira**
E-mail: **rosanesilveira@hotmail.com**

10) CEPSH – UFSC E RESOLUÇÃO 510/6

De acordo com o trecho disponível no site da CEPSCH (Comitê de Ética em Pesquisa com Seres Humanos), o comitê "é um órgão colegiado interdisciplinar, deliberativo, consultivo e educativo, vinculado à Universidade Federal de Santa Catarina, mas independente na tomada de decisões, criado para defender os interesses dentro de padrões éticos". A CEPSH-UFSC se encontra no *Prédio Reitoria II, 4º andar, sala 401, localizado na Rua Desembargador Vitor Lima, nº 222, Trindade, Florianópolis.* Telefone para contato: *3721 – 6094.*
Além disso, declaro que cumprirei a conduzir a pesquisa de acordo com o que preconiza a Resolução 510/16, que dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais, e se encontra no site da CEPSH – UFSC (http://cep.ufsc.br/).

Ao assinar este documento de assentimento esclarecido e livre, você está aceitando em participar da pesquisa. Duas vias deste documento estão sendo rubricadas e assinadas por você e pelas pesquisadoras responsáveis. Guarde cuidadosamente a sua via, pois é um documento que traz importantes informações de contato e garante os seus direitos como participante da pesquisa. Muito obrigada pela leitura e autorização.

Eu, ____, declaro que li este documento e obtive dos pesquisadores todas as informações que julguei necessárias para me sentir esclarecido e livre em participar da pesquisa *The effect of perception training with synthetic and natural stimuli on BP learners' ability to discriminate the English vowels /æ-ɛ/.*

O meu contato de telefone é ______________________________. O meu e-mail para receber o relatório final é ____.

____
Participante da Pesquisa CPF

____
Elisabeth Ann Bunch Oliveira da Rosa Rosane Silveira
Pesquisadora **Orientadora**

APPENDIX B - BACKGROUND QUESTIONNAIRE

1.Data: ______ / ______ / ______

2.Nome completo: ____

3.E-mail: ____

4.Idade: __________

5.Data de nascimento: _______________

6.Lugar de nascimento: ____

7.Você tem alguma deficiência visual? *(marque a caixinha correspondente)*

•Sim

•Não

8.Você tem alguma deficiência auditiva? *(marque a caixinha correspondente)*

•Sim

•Não

9.Qual é seu nível atual de inglês? *(marque a caixinha correspondente)*

•Iniciante

•Intermediário

•Avançado

10.Por quanto tempo você tem estudado inglês? ________________________

11.Como você tem estudado inglês? *(marque todos que se apliquem):*

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

12.Você passou mais de um mês em alguma cidade fora de Florianópolis?

•Sim*(Se sim, detalhe aqui na 12a por favor)*

•Não*(Se não, vá à pergunta #13)*

12a. Responda sobre as outras cidades onde você passou mais de um mês:

Cidade e país ____________________ Quanto tempo: ___________

Cidade e país ____________________ Quanto tempo: ___________

Cidade e país ____________________ Quanto tempo: ___________

Cidade e país ____________________ Quanto tempo: ___________

Cidade e país ____________________ Quanto tempo: ___________

Cidade e país ____________________ Quanto tempo: ___________

Cidade e país ____________________ Quanto tempo: ___________

Cidade e país ____________________ Quanto tempo: ___________

13.Você atualmente estuda alguma outra língua além de inglês?

•Sim*(Se sim, detalhe aqui embaixo por favor)*

•Não*(Se não, vá à pergunta #14)*

13a. Responda sobre as outras línguas além de inglês que você estuda atualmente:

Língua (atual) #1 ____________________________

Nível atual:

•Iniciante

•Intermediário

•Avançado

Como você estuda esta língua?

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

Quantas horas por semana você estuda esta língua?: _________________

Língua (atual) #2 ____________________________

Nível atual:

•Iniciante

•Intermediário

•Avançado

Como você estuda esta língua?

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

Quantas horas por semana você estuda esta língua?: ________________

Língua (atual) #3 ___________________________

Nível atual:

•Iniciante

•Intermediário

•Avançado

Como você estuda esta língua?

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

Quantas horas por semana você estuda esta língua?: ________________

Língua (atual) #4 ___________________________

Nível atual:

•Iniciante

•Intermediário

•Avançado

Como você estuda esta língua?

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

Quantas horas por semana você estuda esta língua?: ________________

14.Você já estudou uma outra língua além do inglês no passado?

•Sim*(Se sim, detalhe aqui embaixo por favor)*

•Não*(Se não, vá à pergunta #15)*

14a. Responda sobre as línguas que você já estudou no passado por favor:

Língua (passado) #1: ___________________________

Quantos anos você tinha quando começou a estudar esta língua? _______

Por quanto tempo você estudou esta língua? ________________________

Nível mais alto em que você chegou:

•Iniciante

•Intermediário

•Avançado

Como você estudou esta língua?

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

Língua (passado) #2: ____________________________

Quantos anos você tinha quando começou a estudar esta língua? _______

Por quanto tempo você estudou esta língua? ________________________

Nível mais alto em que você chegou:

•Iniciante

•Intermediário

•Avançado

Como você estudou esta língua?

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

Língua (passado) #3: _____________________________

Quantos anos você tinha quando começou a estudar esta língua? _______

Por quanto tempo você estudou esta língua? _________________________

Nível mais alto em que você chegou:

•Iniciante

•Intermediário

•Avançado

Como você estudou esta língua?

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

Língua (passado) #4: ___________________________

Quantos anos você tinha quando começou a estudar esta língua? _______

Por quanto tempo você estudou esta língua? _______________________

Nível mais alto em que você chegou:

•Iniciante

•Intermediário

•Avançado

Como você estudou esta língua?

•Na escola

•No colégio

•Na Universidade

•No programa Extracurricular na UFSC

•Escola particular

•Com um(a) professor(a) particular

•Sozinho (usando a internet, livros em casa, etc.)

•Outro: ____

15.Circule o número que corresponde a sua compreensão da(s) língua(s) falada(s) que você estuda ou estudou. *(0 = não entende nada; 7 = entende tudo)*

Língua: ____________________1234567

Língua: ____________________1234567

Língua: ____________________1234567

Língua: ____________________1234567

Língua: ____________________1234567

Língua: _____________________1234567

Língua: _____________________1234567

Língua: _____________________1234567

APPENDIX C - STIMULI VOWEL DURATION AND FORMANT FREQUENCY VALUES

Males /ɛ/ Vowel

Context	Vowel duration	F1	F2	F3
Males - Average English /ɛ/ (HILLENBRAND ET. AL, 1995)	*189 ms*	*580 Hz*	*1799 Hz*	*2605 Hz*
Males - Average BP /ɛ/ (ESCUDERO ET. AL, 2009)	*123 ms*	*518 Hz*	*1831 Hz*	*2722 Hz*
/ɛ/ in carrier word "head" (S2)	182 ms	537.54 Hz	1741.97 Hz	2596.68 Hz
/ɛ/ in carrier word "said" (S4)	231 ms	633.80 Hz	1797.51 Hz	2614.52 Hz
/ɛ/ in carrier word "vet" (S6)	189 ms	606.38 Hz	1746.37 Hz	2311.70 Hz
/ɛ/ in carrier word "bet" (S8)	221 ms	489.76 Hz	1678.23 Hz	2468.79 Hz
Mean Score of four recordings	205.75 ms	566.87 Hz	1741.02 Hz	2497.92 Hz
Standard Deviation (SD)	20.70	56.67	42.32	121.32
Coefficient of Variation (CV)	0.10	0.10	0.02	0.05

Females /ɛ/ Vowel

Context	Vowel duration	F1	F2	F3
Females - Average English /ɛ/ (HILLENBRAND ET. AL, 1995)	*254 ms*	*731 Hz*	*2058 Hz*	*2979 Hz*
Females - Average BP /ɛ/ (ESCUDERO ET. AL, 2009)	*141 ms*	*646 Hz*	*2271 Hz*	*2897 Hz*
/ɛ/ in carrier word "bed" (S1)	230 ms	635.42 Hz	1842.15 Hz	2267.77 Hz
/ɛ/ in carrier word "set" (S5)	151 ms	833.13 Hz	1898.64 Hz	2479.60 Hz
/ɛ/ in carrier word "met" (S7)	218 ms	766.33 Hz	1831.83 Hz	2474.33 Hz
Mean Score of three recordings	199.67 ms	744.96 Hz	1857.54 Hz	2407.23 Hz
Standard Deviation (SD)	34.76	82.12	29.37	98.64
Coefficient of Variation (CV)	0.17	0.11	0.02	0.04

Males /æ/ Vowel

Context	Vowel duration	F1	F2	F3
Males - Average English /æ/ (HILLENBRAND ET. AL, 1995)	*278 ms*	*588 Hz*	*1952 Hz*	*2601 Hz*

Context	Vowel duration	F1	F2	F3
/æ/ in carrier word "had" (S2)	318 ms	713.28 Hz	1628.77 Hz	2347.81 Hz
/æ/ in carrier word "sad" (S4)	300 ms	800.79 Hz	1703.31 Hz	2267.36 Hz
/æ/ in carrier word "vat" (S6)	224 ms	737.83 Hz	1687.18 Hz	2367.25 Hz
/æ/ in carrier word "bat" (S8)	281 ms	689.49 Hz	1637.86 Hz	2432.15 Hz
Mean Score of four recordings	280 ms	735.35 Hz	1665.28 Hz	2353.64 Hz
Standard Deviation (SD)	35.28	41.47	31.65	58.79
Coefficient of Variation (CV)	0.13	0.06	0.02	0.25

Females /æ/ Vowel

Context	Vowel duration	F1	F2	F3
Females - Average English /æ/ (HILLENBRAND ET. AL, 1995)	*332 ms*	*669 Hz*	*2349 Hz*	*2972 Hz*
/æ/ in carrier word "bad" (S1)	259 ms	824.71 Hz	1895.47 Hz	2222.11 Hz
/æ/ in carrier word "sat" (S5)	180 ms	1039.93 Hz	1683.09 Hz	2465.56 Hz
/æ/ in carrier word "mat" (S7)	345 ms	889.01 Hz	1565.47 Hz	2393.55 Hz
Mean Score of three recordings	261.33 ms	917.88 Hz	1714.68 Hz	2360.41 Hz
Standard Deviation (SD)	67.38	90.20	136.56	102.11
Coefficient of Variation (CV)	0.26	0.10	0.08	0.04

Males /i/ Vowel

Context	Vowel Duration	F1	F2	F3
/i/ in carrier word "heed" (S2)	205 ms	335.65 Hz	2237.01 Hz	2827.67 Hz
/i/ in carrier word "seed" (S4)	146 ms	360.55 Hz	1927.85 Hz	2737.92 Hz
/i/ in carrier word "Veet" (S6)	194 ms	322.85 Hz	2340.28 Hz	3060.89 Hz
/i/ in carrier word "beat" (S8)	210 ms	283.55 Hz	2036.76 Hz	2387.54 Hz
Mean Score of four recordings	188.75 ms	325.65 Hz	2135.48 Hz	2753.51 Hz
Standard Deviation (SD)	25.35	27.83	162.10	241.95
Coefficient of Variation (CV)	0.13	0.09	0.08	0.09

Females /i/ Vowel

Context	Vowel Duration	F1	F2	F3
/i/ in carrier word "bead" (S1)	269 ms	377.82 Hz	2515.46 Hz	3218.43 Hz
/i/ in carrier word "seat" (S5)	117 ms	362.00 Hz	2689.19 Hz	3107.52 Hz
/i/ in carrier word "meet" (S7)	357 ms	288.08 Hz	2971.61 Hz	3301.41 Hz
Mean Score of three recordings	247.67 ms	342.63 Hz	2725.42 Hz	3209.12 Hz
Standard Deviation (SD)	99.13	39.11	187.98	79.43
Coefficient of Variation (CV)	0.40	0.11	0.07	0.02

APPENDIX D - SHAPIRO-WILK NORMALITY TEST RESULTS

Data for research question 1

	Statistic	df	Sig.
Pre-Test /ɛ/			
Control	.863	19	.011
Natural	.912	22	.053
Synthetic	.913	15	.152

Post-Test Familiar /ɛ/			
Control	.846	19	.006
Natural	.914	22	.056
Synthetic	.938	15	.359

Pre-Test /æ/			
Control	.912	19	.082
Natural	.939	22	.192
Synthetic	.896	15	.082

Post-Test Familiar /æ/			
Control	.933	19	.195
Natural	.882	22	.013
Synthetic	.885	15	.057

Data for research question 2

	Statistic	df	Sig.
Pre-Test /ɛ/			
Control	.863	19	.011
Natural	.912	22	.053
Synthetic	.913	15	.152

Post-Test Unfamiliar/ɛ/			
Control	.923	19	.126
Natural	.928	22	.109
Synthetic	.843	15	.014

Pre-Test /æ/			
Control	.912	19	.082
Natural	.939	22	.192
Synthetic	.896	15	.082

Post-Test Unfamiliar /æ/			
Control	.948	19	.363
Natural	.865	22	.006
Synthetic	.908	15	.126

Data for research question 3

	Statistic	df	Sig.
Pre-Test /ɛ/			
Control	.863	19	.011
Natural	.912	22	.053
Synthetic	.913	15	.152

Post-Test Familiar /ɛ/			
Control	.846	19	.006
Natural	.914	22	.056
Synthetic	.938	15	.359

Delayed Post-Test Familiar /ɛ/			
Control	.923	19	.126
Natural	.890	22	.019
Synthetic	.795	15	.003

Pre-Test /æ/			
Control	.912	19	.082
Natural	.939	22	.192
Synthetic	.896	15	.082

Post-Test Familiar /æ/			
Control	.933	19	.195
Natural	.882	22	.013
Synthetic	.885	15	.057

Delayed Post-Test Familiar /æ/			
Control	.936	19	.221
Natural	.910	22	.046
Synthetic	.861	15	.025

Data for research question 4

	Statistic	df	Sig.
Vocabulary Score			
Control	.935	19	.214
Natural	.948	22	.284
Synthetic	.971	15	.869