UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE ESTUDOS DA LINGUAGEM

FILIPE MODESTO

Acoustic analysis of lexical stress in English by Brazilian Portuguese speakers, and inferences of production and perception

Análise acústica do acento lexical em inglês por falantes do português brasileiro e inferências quanto a produção e percepção

CAMPINAS
2019

FILIPE MODESTO

Acoustic analysis of lexical stress in English by Brazilian Portuguese speakers, and inferences of production and perception

Análise acústica do acento lexical em inglês por falantes do português brasileiro e inferências quanto a produção e percepção

Dissertation presented to the Institute of Language Studies of the University of Campinas in partial fulfillment of the requirements for the degree of Master, in the area of Linguistics.

Dissertação de mestrado apresentada ao Instituto de Estudos da Linguagem da Universidade Estadual de Campinas para obtenção do título de Mestre em Linguística.

Adviser: Prof. PhD. Plínio Almeida Barbosa

Este arquivo digital corresponde à versão final da dissertação de mestrado defendida pelo aluno Filipe Modesto e orientada pelo Prof. PhD. Plínio Almeida Barbosa.

CAMPINAS
2019

Agência(s) de fomento e nº(s) de processo(s): Não se aplica.

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Estudos da Linguagem
Crisllene Queiroz Custódio - CRB 8/8624

M72a Modesto, Filipe, 1991-
Acoustic analysis of lexical stress in English by Brazilian Portuguese speakers, and inferences from production and percetion / Filipe Modesto. – Campinas, SP : [s.n.], 2019.

Orientador: Plinio Almeida Barbosa.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.

1. Fonética acústica. 2. Língua inglesa - Estudo e ensino - Falantes de português. 3. Língua inglesa - Análise prosódica. 4. Língua inglesa - Acentos e acentuação. 5. Análise prosódica (Linguística). I. Barbosa, Plinio Almeida, 1966-. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Análise acústica do acento lexical em inglês por falantes do português brasileiro e inferências quanto a produção e percepção
Palavras-chave em inglês:
Acoustic phonetics
English language - Study and teaching - Portuguese speakers
English language - Prosodic analysis
English language - Accents and accentuation
Prosodic analysis (Linguistics)
Área de concentração: Linguística
Titulação: Mestre em Linguística
Banca examinadora:
Plinio Almeida Barbosa [Orientador]
Sandra Madureira
Leônidas José da Silva Junior
Data de defesa: 08-03-2019
Programa de Pós-Graduação: Linguística

BANCA EXAMINADORA:

Plínio Almeida Barbosa

Sandra Madureira

Leônidas José da Silva Junior

IEL/UNICAMP
2019

Ata da defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós Graduação do IEL.

"If you can dream it, you can do it"

Walt Disney

To mother,

because blood is thicker than water.

Acknowledgements

No one can truly succeed alone. My support network made me gather the strength and resilience necessary to get me this far in life, and today I reach another important milestone because of them. Let me now give credit where credit is due.

First, to my advisor, professor Plinio Barbosa who with such kindness and patience was able to trust me and push me forward to pursue this degree, even in the moments when I myself didn't think I could. His technical brilliance and sophisticated passion for knowledge are inspirational and unique.

The University of Campinas (UNICAMP) and the Institute of Language studies, for having amazing and capable faculty and for opening opportunities to me and many other students who want to dedicate themselves to academia, with distinct accreditation. I hope this can keep up for future generations.

To Ithaca College and the faculty at the Department of Speech Language Pathology and Audiology and the Department of Writing for taking me in in all my endeavors since college, whenever I reached out for them, never doubting my capacity. Thank you, professors Marie Sanford, Louise Cannon, Tina Caswell, Jana Waller, Amy Rominger, and Luanne Andersson.

To CAPES, for providing the financial resources to fund my studies.

To my parents Josiani and Sérgio Modesto. My foundation. The ones who devoted love, time, money, patience, and much more in order to get me through college, and now grad school. I could not have done any of this without your support and love. I hope my attempts to make you proud were successful.

Last, but never least, to my beloved friends, quintessential to give me energy and motivation to enter the program, and most of all, finish it. You all have an important contribution in this process. Some of the names that couldn't be out of this list: Hipassia (aka pipa) for believing there is always a way to work things out; Stephen Gray, who proved that friendship and distance can coexist and who has always been an inspiration for me professionally and personally, and a truly believer that I can go farther; Lolly, my former advisor and special friend who taught how to roll up my sleeves and get to work with love and determination since I was a mere freshman; Professor Filomena, who was the first to tell me I had a special trait for linguistics; My dears: Amy, Dom, Zilda, Sandra B, Sandra L, Sonia, Nivia, Bia. I am sorry for the ones I could not remember at the time to add to this list, but if you were meaningful, you certainly feel it.

Resumo

As produções orais de adultos em L2 são caracterizadas por algum grau de sotaque estrangeiro. Dentre as diversas características prosódicas, contribui para esse sotaque a não realização adequada do acento tônico (AT), que tem papel preponderante para a estruturação prosódica da fala. O português brasileiro (PB) possui padrão paroxítono predominante, enquanto o inglês americano (IA) possui acento inicial predominante. Essas divergências na natureza do AT nas duas línguas podem levar a erros na produção de falantes do PB no IA que vão além das distorções segmentais. 29 falantes do PB de quatro níveis de proficiência de IA participaram de testes de produção e percepção de AT. O corpus utilizado (BRENGLISH-STRESS) controlou para posição acentual e cognatismo da palavra. Os dados dos parâmetros acústicos de produção dos participantes, assim como os escores na marcação da posição acentual foram coletados e comparados com o controle de sujeitos nativos do IA. Os áudios dos falantes brasileiros também foram submetidos à análises de grau de sotaque pelos sujeitos nativos. O presente trabalha visa então estudar como se dá a realização do AT em inglês por falantes brasileiros de diversos níveis, buscando observar se há diferenças tanto na colocação correta do acento, quanto nas características acústicas das produções entre os níveis. Objetivou-se também tentar estabelecer uma relação entre produção e percepção do AT, tendo o IA como L2. Duração silábica e intensidade relativa foram os parâmetros mais utilizados tanto para falantes do PB quanto IA para a realização do AT. O cognatismo das palavras influenciou apenas a posição de acentuação e não a realização acústica das palavras. A percepção do AT foi mais fácil do que a percepção para todos os níveis, embora o mesmo não seja verdade quanto a discriminação entre níveis de acento primário e secundário.

Palavras-chave: sotaque estrangeiro; fonética acústica; prosódia.

Abstract

Oral productions of adults in L2 are characterized by at least some degree of foreign accent. Amidst the many prosodic features, the mispronunciation of lexical stress (LS) contributes for the perception of a foreign accent by native speakers, since LS is fundamental for the prosodic structure of speech. Brazilian Portuguese (BP) stress pattern is more paroxytone, while American English (AE) has more words with initial stress. These differences in LS between both languages can lead to mistakes in the productions of BP speakers in AE. 29 BP speakers of four different levels of proficiency participated in tests of production and perception of LS. The corpus (BRENGLISH-STRESS) used in this study controlled for stress position and cognate status of the words. The acoustic data of the participants' productions, as well as the scores in assigning stress to the correct position were collected and compared to the American controls. The audios of the BP speakers underwent accent ratings by the AE speakers. This work aims to study how LS is realized in English by BP speakers of different levels, analyzing whether there are differences in the assignment of correct word stress, as well as the acoustic features of the productions throughout levels of proficiency in the language. Another aim was to try to establish a relationship between production and perception of LS, having AE as an L2. Syllable duration and relative intensity were the parameters used by both BP and AE speakers to realize lexical stress. The cognate status influenced the position of stress rather than the acoustic realization of it. LS perception was easier than its production for all levels of proficiency, although discrimination between primary and secondary levels of stress was not.

Keywords: foreign accent; acoustic phonetics; prosody.

List of Figures

List of Tables

List of Abbreviations

Acronym	Meaning
AE	American English
aka	"as known as"
ANOVA	Analysis of Variance
AOL	Age of learning
AP	Acoustic Parameter
apud	"in the writings of"
AR	Accent rating
BP	Brazilian Portuguese
C	Cognates
CEF	Common European Framework
CP	Correct position
DMRT	Duncan's Multiple Range Test
DUR	Duration
EE	Exam English
EH	Early (learners of L2) High (use of L1)
EL	Early learners (of an L2)
EL	Early (learners of L2) Low (use of L1)
ES	Effect Size
F0MED	Median of F0
F0PEAK	Peak of F0
F0SD	Standard deviation of F0
FA	Foreign accent
FP	Favorite position
L1	First language
L2	Second language
LH	Late (learners of L2) High (use of L1)
LH	Late (learners of L2) High (use of L1)
LL	Late learners (of an L2)
LL	Late (learners of L2) Low (use of L1)
LOR	Length of residence
LS	Lexical stress
N1	Proficiency level 1
N2	Proficiency level 2
N3	Proficiency level 3
N4	Proficiency level 4
NC	Non cognates
NE	Native English
P1	Position 1 (of stress)
P2	Position 2 (of stress)
P3	Position 3 (of stress)
PL	Proficiency level
RELINT	Relative intensity

SNR	Signal to Noise ratio
TOTINT	Total intensity

Contents

0 Motivation

Maybe one of the most fascinating things early learners of English as an L2 notice (having Brazilian Portuguese as L1) is how the vowels glide from one to the other creating a melodic pattern for the language. Listening closely to native speakers is a very helpful technique where one can understand how sounds are pronounced in different situations. How "ea" can sound so differently across the words bear, steak, and meat. The segmental features are very tactile, easy to be perceived and differed from one another in my opinion. That is when I decided to explore what was above the segments (literally) – the suprasegmentals.

It wasn't until I noticed something else was very different from Portuguese and English, something that went beyond the difference between the pronunciation of two segments but how the same segment can be pronounced differently, that I started to want to do research within prosodic phonetics. Prosody was the key for what used to bother me about the differences in my English and a native speaker. I felt like my syllables could be pronounced native like but I was still missing something. It was not about the structure of the segments, it was about how they talk to each other, how they alternate and create rhythm. After these insights, my pronunciation improved a lot and I started to realize how different the two languages that always surrounded me were. I needed to understand more, I needed to see if my fellow Brazilians felt and reacted to pronouncing and perceiving foreign prosody like I did.

In order to understand the concepts approached in this dissertation, some background in foreign accent, prosody, and lexical stress is needed.

1 Foreign Accent

It is believed that children begin the language acquisition process by incorporating rhythmic aspects of a language (Mehler, 1988; Cutler and Mehler, 1993). Such particularity allows them to make production and perception adjustments regarding prosody. For years, it was believed that this ability of acquiring prosodic features becomes less flexible over time along with other linguistic abilities, making learning process of an L2 more complex. Age-related changes in this learning process would be due to loss of neural plasticity. This idea of a critical period (CP)[1] was first introduced by Lenneberg (1967) and the authors who base themselves in it to study L2 basically state that mastery of an L2 is no longer possible after a so-called putative CP. Despite highly controversial, the theory applied to L2 literature claims that:

> *"Neurological maturation might reduce neural plasticity (Penfield 1965; Lenneberg 1967), leading to a diminished ability to add or modify sensorimotor programs for producing sounds in an L2 (Sapon 1952; McLaughlin 1977)." (p.234)*

Some have even suggested several CP over lifespan, affecting different linguistic abilities (FATHMAN, 1975; SELIGER, 1978). Initially, one would not be able to develop native-like pronunciation of an L2 if it was learned after the CP. A looser idea of learning

[1] Different CP ages have been proposed: 12 years old (SCOVEL, 1988) and 15 years old (PATKOWSKI, 1990) being the most known ones. After these ages, learning would be more difficult and a foreign accent in L2 inevitable.

development, called a *sensitive* period proposes a more gradual degree of perceived foreign accent over time, but the terminology is interchangeably used with CP.

In the article by Piske, MacKay, and Flege (2001), the authors thoroughly reviewed the literature in overall degree of L2 accent focusing on the main causes of accent, according to the articles. These variables were: age of learning (AOL), length of residence (LOR) in an L2 speaking area, gender, formal instruction, motivation, language learning aptitude, and L1 use. None of the variables presented consistent evidence of cause, besides AOL and L1 use. AOL would be considered more as "an index of the state of development of the L1 system. The more fully developed the L1 system is when L2 learning commences, the more strongly the L1 will influence the L2" (p. 196). AOL alone is not a good indicator of L2 mastery success. In an experiment from 1997, Flege and colleagues studied two groups of early Italian-English children (mean age= 6 years old) that differed significantly in regards to L1 use. Both groups were found to speak English with a detectable foreign accent.

Going in the opposite direction, studies have also found post lingual adults who were able to master an L2 with native-like competence. For instance, Moyer (1999) showed a native English speaker (age= 22 years old) with a mean rating in German higher than the rating given to one of the four Germans in the control group. Bongaerts et al. (1997) showed that five Dutch learners of English were able to score as much as the native English controls, having begun to learn the language no longer than the age of 16 years old. Therefore, AOL alone cannot be a sole predictor of L2 success, although we do not deny its relevance in the process of learning a new language without as many difficulties as for late learners.

In the same article by Piske, MacKay, and Flege, they also ran an experiment with over 70 Italians who went to Canada, divided in groups according to their AOL. The groups were comprised by Italians who started learning English as children (early bilinguals) and in their adolescence or later (late bilinguals). The speech materials were 270 read sentences rated by nine native speakers of English in a 9-point scale from "No FA [foreign accent]" to "Very Strong FA". Figure 1 below presents the results for the ratings. As we can see, regardless of L1 use, early bilinguals received higher scores than late bilinguals, and within these two groups, the higher scores were always for those who used less of their L1.

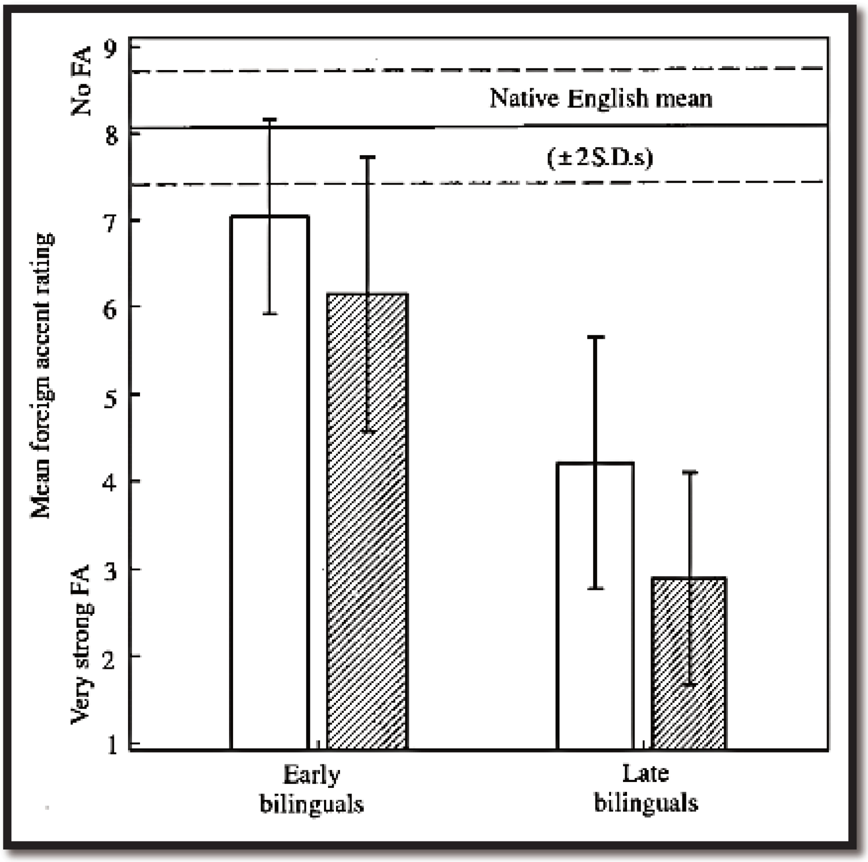


Figure 1: mean ratings by Canadian listeners for the Italian learners of English in the Study of Piske, MacKay, and Flege (2001). White bars= low L1 use, Hashed bars: high L1 use.

Analyses were carried out for all four indicators together and all the possible combinations of parameters being partialled out. It is interesting to notice that correlation between foreign accent and AOL was so solid that, even with each of the other variables partialled out one at a time, and all three of them partialled out altogether, the correlation kept significant! The same thing happened to the amount of L1 use, showing that AOL and L1 use are independent predictors of the Italian subject's degree of foreign accent in English. Overall, the results of the study corroborate with the view that "ultimate attainment in the pronunciation of an L2 is dependent on various factors, not just on the state of neurological development at the age of first intensive exposure to the L2" (p. 212).

Many years later, Flege (2018) stated once and for all that foreign accents are indeed related to AOL, but refutes the CP theory based on the success of late learners throughout several studies. In his very emphatic article entitled "It's input that matters most, not age", he mentions his work with MacKay (2004). In the study, 2 groups of Italian speakers of English that differed orthogonally in age of arrival (8 x 20 years old) and percentage of Italian use (8% x 48%) were tested for the perception of [i] and [I] in English, because it is known that Italians hear these two vowels interchangeably. Figure 2a presents the categorial discrimination of the vocalic pair, while 2b presents the rates of error detections of the vowels in words where they were misplaced on purpose. Only late learners with high use of Italian (LH) were significantly different from native English (NE) speakers. In Figure 2b, the lack of difference between NE and EL, and EL and LH are "incompatible with the CP hypothesis" (p. 920).

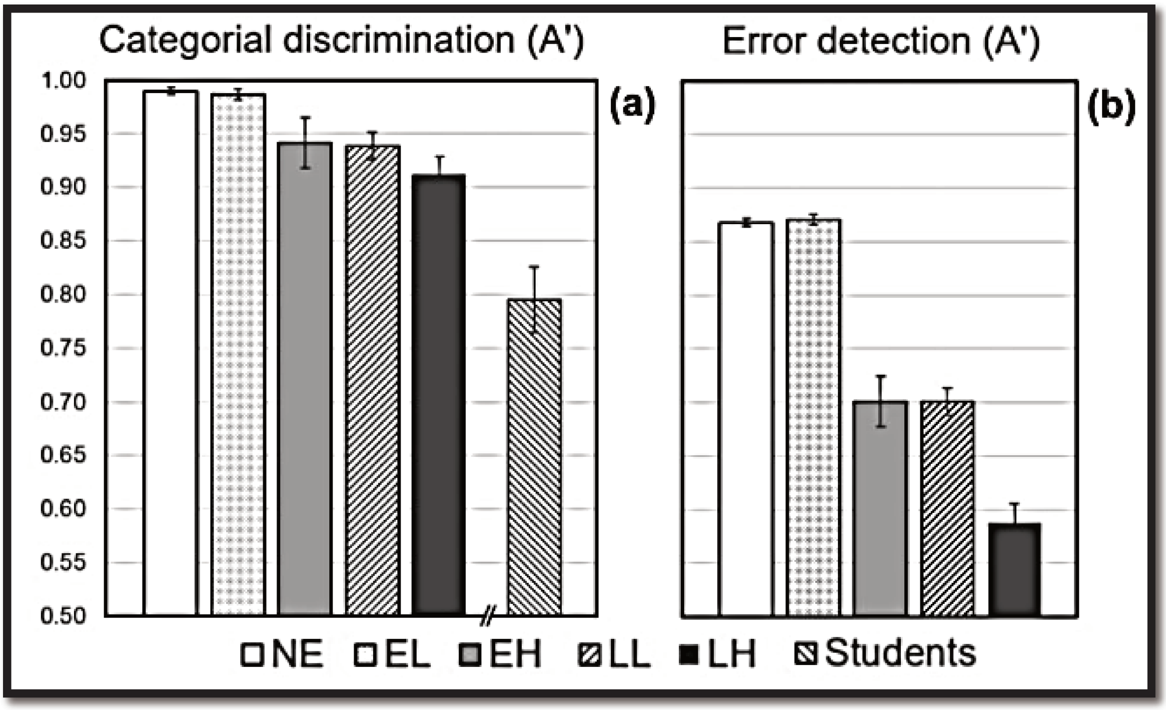


Figure 2: (a) Discrimination and (b) error detection of English /i/ and /ɪ/ by groups of Italians differing orthogonally in age of arrival in Canada (Early vs. Late) and amount of Italian use (Low vs High). NE= Native English, EL= Early Low, EH= Early High, LL=Late Low, LH= Late High

Foreign accents in English are common in non-native speakers of the language. Accent occurs when listeners of a language detect divergences from English phonetic norms along a wide range of segmental and suprasegmental dimensions (Flege, 1995).

Regardless of which theory addresses the issue of foreign accent, the most referred features of an accent rely on a segmental level. Many studies have shown that non-native speakers must listen the same word many times in order to fully recognize it, representing a greater difficulty at the level of a segment (Miyawaki et al. 1975; Flege and Eefting 1986; Flege and Hillenbrand 1986, 1987). Flege (1995; p.239) brings to our attention 7 hypotheses regarding second language acquisition that should be considered when studying foreign accent:

- H1: sounds in L1 and L2 are related to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level;
- H2: if a sound from L2 that differs from the closest sound in L1 can be perceived with at least some phonetic differences, a new phonetic category can be established for the L2 sounds.
- H3: the more different two sounds are perceived between an L2 sound and the closest sound from L1, the more likely it is that phonetic differences between them will be distinguished.
- H4: the likelihood of phonetic differences between L1/L2 sounds, and between L2 sounds that are non-contrastive in the L1 being distinguished decreases as the age of learning increases.
- H5: category formation for an L2 sound may be blocked by the mechanism of equivalence classification. If that happens, a single phonetic category will be used to process linked sounds between L1-L2 perceived the same. Eventually, the diaphones will resemble one another in production.
- H6: the phonetic categories established for L2 sounds by a bilingual may differ from a monolingual's if:
    1. The bilingual's category is separated from an L1 category to maintain phonetic contrast in a common L1-L2 phonological space.
    2. The bilingual's representation is based on different features, or feature weights than a monolingual's.
- H7: The production of a sound eventually corresponds to the properties represented in its phonetic category representation.

To better illustrate how the phonetic features are related, Figure 3 shows how sound categories can be mentally represented.

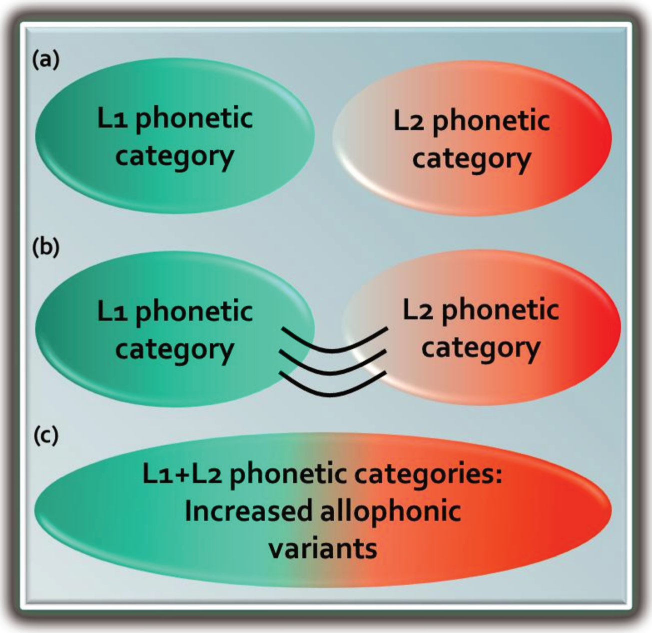


Figure 3: mental representation of phonetic categories

In the situations a) and b), there are two different categories representing the phonetic groups of both L1 and L2, whereas, in the latter, we observe an interlinguistic interaction between the two representations, producing the foreign accent. An example that elucidates this situation is the pronunciation of the word *two*. Many Brazilians, instead of producing the aspirated, voiceless alveolar stop [t^hu], distort its production by yielding a voiceless palatal affricate [ʧu], which makes the word *two* sound like the word *chew*. As we can see in c), instead of two different categories, we have one category with expanded allophonic variants. The crossover between the two phonetic categories showed in b) does not happen randomly. It is predetermined according to the nature of the subject's L1.

2 Prosody and Accent

Most studies discuss foreign accent in a segmental level or trying to associate foreign accent with segmental factors (Piske, MacKay, Flege, 2001; Lenneberg, 1967; Asher & García, 1969; Thompson, 1991; Suter, 1976; Tahta et al, 1981; Purcell & Suter, 1980) investigating either vowels or consonants, leaving behind prosodic aspects.

Prosodic features have clearly been studied less than segments in regards to foreign accent, even though errors in prosody production have a greater contribution to perceived foreign accent than does syllable structure and segmentals (ANDERSON-HSIEH, JOHNSON, KOEHLER, 1992; JILKA, 2000). In this particular study by Anderson-Hsieh et al, 11 different languages were studied, and prosody was the only set of parameters always perceived as different by the 60 native speakers of English who were rating the accents. In a study by Munro (1995), filtered and unfiltered speech of native Mandarin speakers in English was rated by native American speakers. In both scenarios, the ratings for foreign accent were higher than for the native speakers of English used as control, showing that regardless of segmental information, divergences in prosody alone can be an indicator of foreign accent. The studies that prioritize prosody, mostly focus on intonation, even though there are a few studies in pitch range and stress (KANG, 2010) and speech rate (MUNRO, DERWING, 2001), for instance.

In another study by Jilka (2000), the relevance of prosody was also pointed out. Different sentences produced by Germans who were learning English and vice versa were used in perception experiments, where the audio samples underwent low-pass filtering and the raters had to identify the language and accent of the speaker. The raters from the study were able to identify the speakers' languages using prosodic information alone.

According to the discussion so far, it is agreeable that alterations in the production of prosodic features may lead to a higher or lower degree of foreign accent. Although, the relevance of each parameter associated to the given perceived accent vary according to the level of proficiency of the L2 speaker, and even the proximity between L1-L2.

3 Lexical Stress and Accent

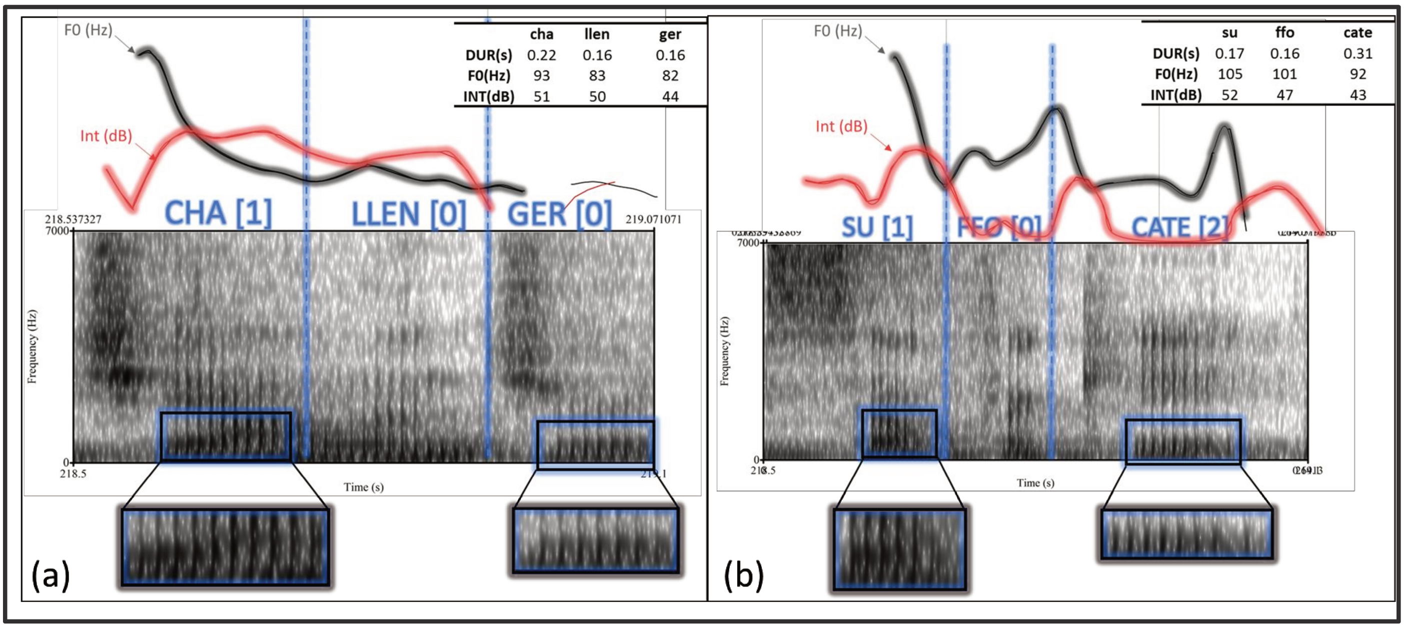
Lexical stress relates to the prominence of a syllable within a word. It is able to affect rhythmic, positional, quantitative, and morphological patterns According to Kager (1995):

> *Prominent syllables are potentially capable of bearing pitch movements with a strong perceptual load. They also tend to be of longer duration, as well as of higher intensity, but both of the latter factors are usually subordinated to pitch. (p.367)*

The author makes some considerations about word stress: every content word has at least one stressed syllable, it is hierarchic, it is rhythmic in systems where stressed and unstressed syllables alternate, and it tends to be enhanced segmentally by vowel lengthening or gemination.

In every language where stress is a contrastive feature, there is going to be at least one syllable primarily stressed within a word. Stress is binary if there are only two levels: stressed and unstressed, and many languages also have a secondary level of stress. Figure 4 below show different spectrograms of words bearing different levels of stress in English, produced by an adult native man in the language. The words were in condition of linguistic focus. In the picture we can see how the acoustic parameters articulate in order to make certain chunk of the word prominent (the stressed syllable).

Figure 4: spectrogram, pitch, and intensity contours of (a) a word with primary stress



in the first syllable and following unstressed syllables (*CHAllenger*), and (b) a word with primary stress in the first syllable and secondary stress in the last syllable (*SUffocate*).

We can see that the pitch contour in 4a has its peak in the stressed syllable and it remains shallow through the following ones, and intensity is high in the first two syllables, and drops in the last one. In the word with secondary stress (Figure 4b), we can see alternating F0 peaks but the highest value of F0 is placed in the primarily stressed syllable. The last syllable in this word is longer than the first - that bears higher level of prominence - due to the effect of final lengthening. Syllables bearing primary stress are often longer than all the other ones. We can observe in both parts of the picture the change in vowel quality between syllables in different conditions of stress (stressed-unstressed in 4a, and 1ry stress-2ry stress in 4b), observing the F1-F2 pattern in the zoomed grey scale on the spectrograms. The first and second formants of the vowels in the most stressed syllable is always darker, indicating more energy to that particular phone.

From the picture, it is interesting to notice that even though they both had syllables with primary stress, they were acoustically realized in different ways: in 4a, an increase of all three prosodic parameters, while in 4b the stressed syllable was shorter but still had higher F0 and intensity values. These differences in the acoustic realization of stress occur cross-linguistically as well. Despite being both stress languages, Brazilian Portuguese (BP) and American English (AE) do not realize stress the same way, nor do they follow the same stress patterns. Ahead, some of such differences between both languages.

4 Lexical Stress in Brazilian Portuguese (BP)

In BP, word stress is distinctive for a set of words pairs, with a special function setting differences across grammatical categories (LEE, 1995). Some words that fully elucidate that are: *sábia (wise person, female), sabia (past tense of "to know", third-person singular),* and *sabiá (thrush bird)*. Each word has a different meaning according to word stress placement (Gonçalves Vianna, 1883; Câmara Jr, 1979; Delgado Martins, 1988; Massini-Cagliari, 1999, Barbosa, 2006). BP words have three possibilities to stress placement. Words can be: oxytone (stress in the last syllable), paroxytone (stress in the penultimate syllable), or proparoxytone (stress in the antepenultimate syllable). Table 1 shows the properties of word stress in BP (Cantoni, 2009).

Table 1: properties of word stress in Brazilian Portuguese.

Placement of Stress	*One of the three last syllables*
Acoustic-articulatory Nature	*Increase in duration and intensity; Vowel contrast between stressed and post-stressed position*
Function	*Lexical contrast, with restrictions*
Stress Pattern distribution	*The paroxytone pattern is the most frequent*

In a study by Barbosa, Eriksson and Akesson (2013), acoustic correlates for lexical stress were monitored for three different speaking styles: informal interview, phrase reading, and word list reading. 10 subjects (5 male, 5 female) participated in the study. The results showed the measures of centrality for F0 and Intensity did not signal lexical stress in BP, since they are both intonation dependent, that is, not lexical exclusively. However, duration, vowel spectral emphasis, and F0 standard deviation in decreasing order of relevance are distinct between stress levels in all three speaking styles.

5 Lexical Stress in American English (AE)

As mentioned before, English is a stress language, but that does not mean stress works the same as in BP. On the contrary, it is different in regards to the pattern of accentuation, as well as in the acoustic realization. Stress in AE has an important role in the derivational morphology of the language (CUTLER, 2015). That is, adding affixes to words may create a different word of different grammatical class, shifting stress with it (e.g.: *adMIRE-admiRAtion, deMOcracy-demoCRAtic*). Even more interesting, when stress shifts are of greater magnitude than segmental changes, this can lead to alterations in meaning (e.g.: *CONtract-conTRACT*, *PERvert-perVERT*).

In AE, a stressed syllable within a word can be longer, louder, and of higher pitch, compared to unstressed syllables. The main features of lexical stress in AE are described in Table 2:

Table 2: features of stressed words in English.

F_0	***Increased fundamental frequency***
Intensity	*Increased intensity*
Duration	*Increased duration*
Vowel quality	*Differentiation of the stressed vowel*

Table 2 shows the increase of fundamental frequency, intensity, and duration that occurs when a syllable is stressed in English. Not placing the frequency peak in the stressed syllable indicates foreign accent in the non-native speaker's production (Zhang, 2008). Figure

5 shows the waveforms and the spectrograms of the word *resume* read with all possible stress patterns for this word.

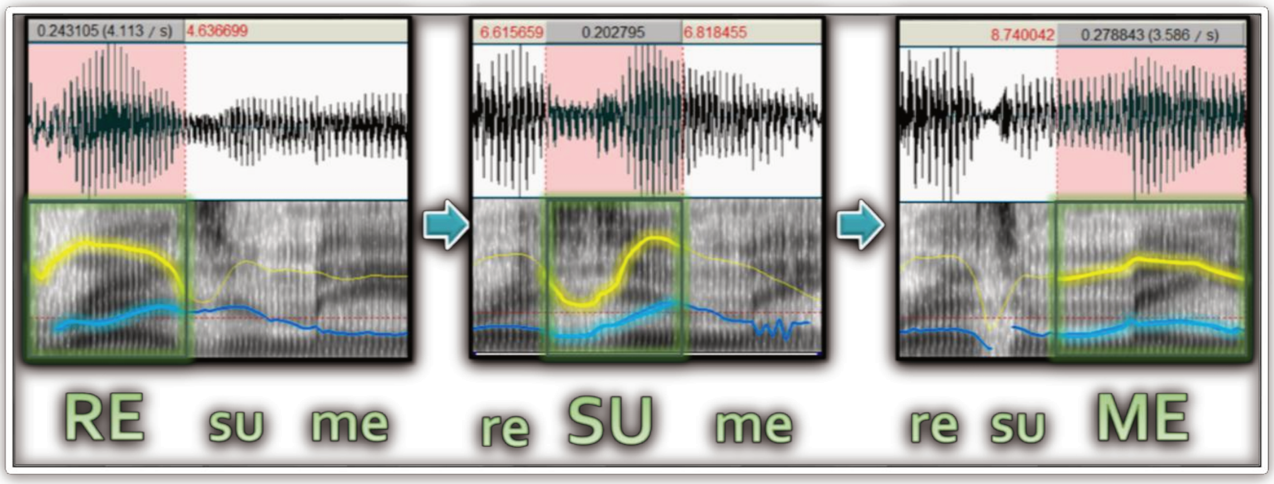


Figure 5: representation of acoustic variation across with stress the word *resume*.

The three instances of this word were recorded in laboratory conditions and the sound file was analyzed using the free software Praat (Boersma and Weenink, 2018). The yellow curve corresponds to intensity, the blue curve corresponds to fundamental frequency, and the horizontal extension of the green window corresponds to the duration of the stressed syllable. The increase of the three acoustic parameters mentioned here within the stressed syllable is apparent. This would suggest that different prosodic patterns are relevant as an indication of foreign accent.

Regarding the stress pattern, Anne Cutler (2015) brings that:

> “there is a highly significant tendency for stress in English words to fall on the initial syllable, and

> this tendency is even greater in real speech samples (Cutler and Carter, 1987). There is an obvious reason for this: about a quarter of the vocabulary consists of words with unstressed initial syllables, but most of the words in this set have a relatively low frequency of occurrence (*pollution, acquire, arithmetic*). The higher frequency words, i.e., the ones most often heard in real speech, are shorter and more likely to have just a single [primarily] stressed syllable (*garbage, borrow, numbers*) or the only syllable (*trash, take, math*)" (p. 110).

Another strategy used in English to mark stress goes to a segmental level. Vowels of stressed syllables must have a full vowel quality, and vowel reduction of pre-stressed or post stressed syllables may occur to weaken the syllable. The most common reduction occurs to the centralized vowel (schwa). Even though vowel quality also differs with stress status in Portuguese, more extreme cases of vowel reduction occurs only in European Portuguese.

6 Methods

6.1 The Corpus

Our corpus (BRENGLISH-STRESS) was comprised by 73 words of the English language (APPENDIX IV). The same words were used in both production and perception tests. The words were divided in two major categories, based on the cognate relationship with BP (see Figure 6 for a representation of the word categories). Within each major category (cognates (C) and non-cognates (NC)), there were four other subcategories: -ate, -tion, and –ct, which indicate that the words presented that given suffix. The C-other and NC-other categories are cognates and non-cognates (respectively) that did not end in any of the mentioned suffixes.

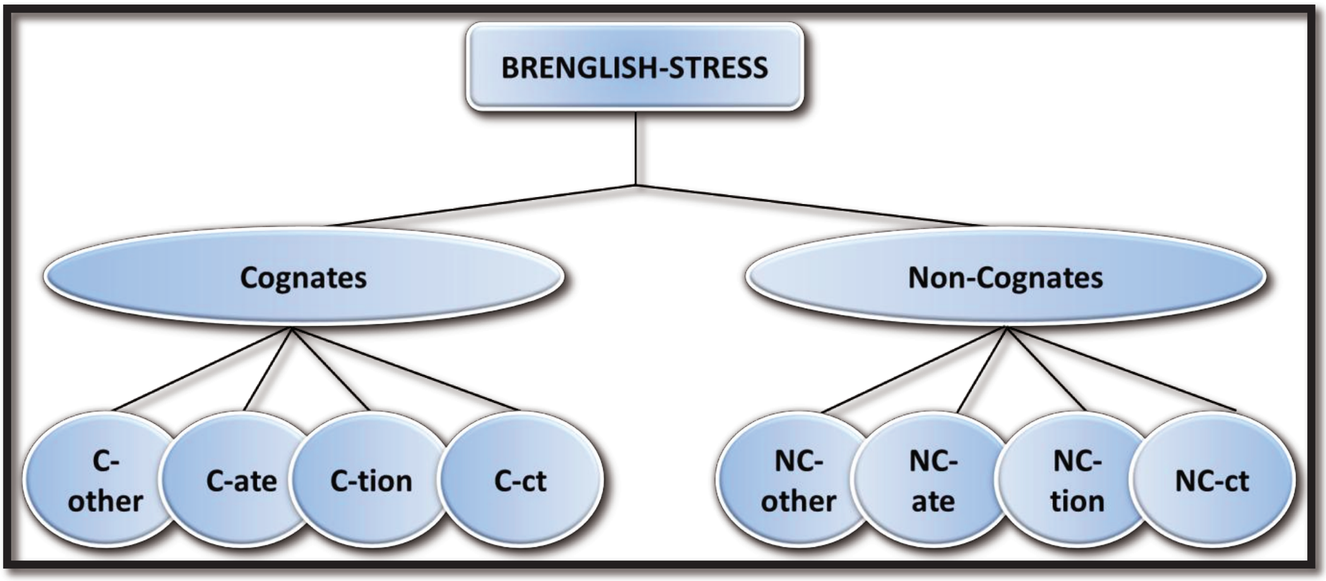


Figure 6: representation of the categories of words used in this study.

We controlled for word length, using only three syllable words. Therefore, there were only three possible primary stress placement positions available, which we here call: P3= first syllable, P2= second syllable and P1= last syllable of the word. As can be noted, the syllable count starts from the end of the word. The stress distribution of words are in Table 3. The P3 words composed most of the corpus, since the majority of the lexical words in English are stressed in the first syllable (as previously discussed in the introduction).

Table 3: distribution (in %) of stress position of the words.

Stress Position		
P3	**P2**	**P1**
51%	29%	20%

About the same proportion of stress position reported in Table 3 was maintained within C and NC subcategories (Figure 7).

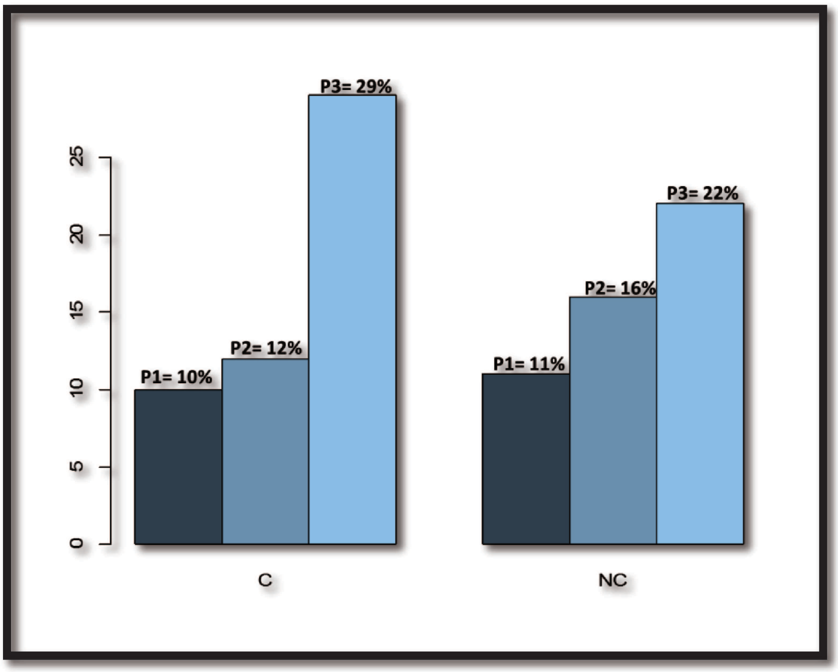


Figure 7: stress position distribution per cognate category.

Out of the 73 words, 37 were C, and 36 NC. In 14 of the C words and 13 of the NC words subcategories suffix was controlled (Table 4). By doing that, we were able to observe if BP was taking part in AE linguistic stress, since the suffixes influence the placement of lexical stress in English (for C and NC words). However, for C words, an interference of L1 is expected to take place a major influence in stress position. For instance, all words ended in –ate are P3 (e.g.: DEmonstrate), words ended in –tion ate P2 (e.g.: auDItion), and words ended in –ct are P1 (e.g.: disinFECT).

Table 4: distribution of number words per category.

	Word Type							
	C-other	**Cate**	**Ction**	**Cct**	**NC-other**	**NCate**	**NCtion**	**NCct**
N	23	5	5	4	23	5	5	3

This stress movement helps us test the hypothesis that the greater the distance difference between the native stress position of a word with the respective cognate in L1, the harder it gets to stress such word properly in the L2 (e.g.: SUffocate [AE] and its correspondent sufoCAR [BP] with extreme opposite stress placements). In case both words coincide in stress, it is presumably intuitive that the error rates will be low (e.g.: disinFECT [AE] and desinfeTAR are both stressed in the last syllable). Figure 8 exemplifies that. The green bars are supposed to express the distance between the stressed syllable in AE and the correspondent word in BP.

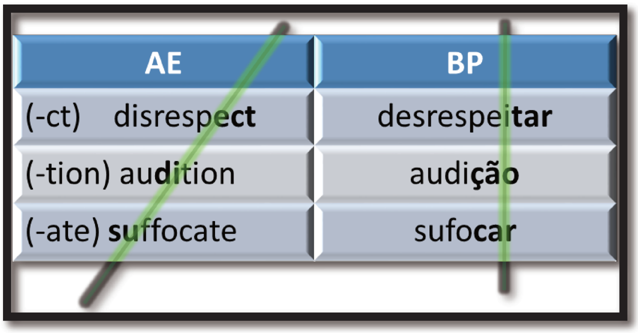


Figure 8: example of stress distance across cognates of BP and AE.

Another important aspect of these words is that 49% have secondary stress. All of these were either P3 or P1. The P2 words have only primary stress.

Regarding the morphology of the words, the corpus was composed by 47% of verbs and 53% of non-verbs, being either a noun (41% of the words) or an adjective (12% of the words). All cognate verbs were also verbs in BP. From all words, 82% were bimorphemic, and the rest was monomorphemic, except that only 3words had three morphemes (*confection*, *employer*, and *misfunction*).

6.2 Experimental Flowchart

The step by step of this study will be presented in detail in the following sections. We also present the experimental flowchart (Figure 9), just so the reader can have a clearer picture about the complete experimental design and the development of each step and its rationale.

As can be seen from Figure 9, the first step is the English level evaluation. It was assessed in three different ways: a self-evaluation that was the simple rate of each participant regarding his/her level of English in a progressive 1-4 scale. The self-evaluation helps us find if there is a relationship of it with the level measured objectively, with the Exam English (EE) test (a standardized test for Vocabulary/Grammar and Listening, presented in section 6.3.1). Another way the level of English was verified was with the accent rating by American subjects (section 6.6.3). The English level was performed in the latest stage of the research, given that the entire corpus had to be recorded first in order to be rated.

Once the screening from the EE step had been completed and the levels properly assigned, subjects who fell in the inclusion criteria of the research were recruited to participate in the production and perception tests. The production test was always the first (section 6.6.1), followed by the perception test (6.6.2). The participant always had to say the first three sentences of the production test as practice, and the words were then randomized again so that the test could really begin. The same measure was taken for the perception test.

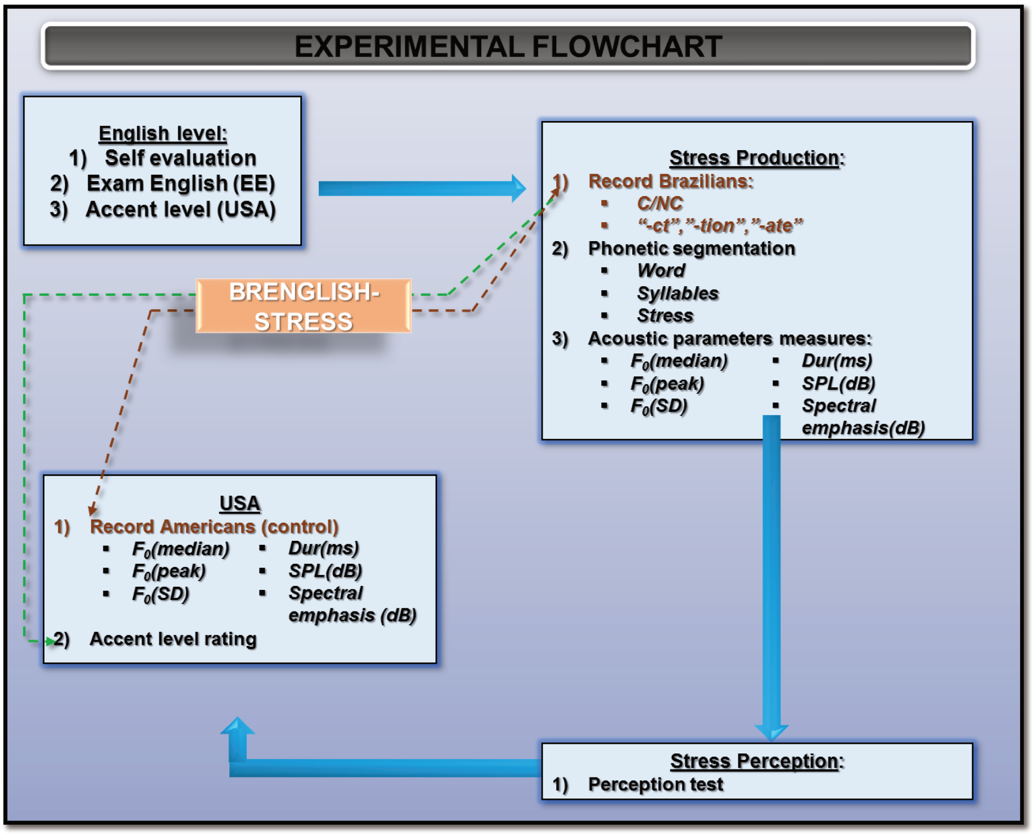


Figure 9: experimental flowchart of the research.

After the Brazilians of all four levels were recorded, samples of whole sentences were extracted from the files and helped compile the accent-rating test. The test was performed with the Americans from the control group. The brown dashed lines in the picture indicates that the same corpus used in the production test for the Brazilians was used for the Americans, and the green dotted line indicates that the audios from the production test by the Brazilians were used to the accent level rating. The third square in the picture (with USA written in it) refers to the part of the research that was conducted overseas, in Ithaca-NY (USA).

6.3 English Level Tests and Group Placement

All subjects took English level tests for Vocabulary/Grammar and Listening tasks. According to their scores, they were then placed in different subgroups according to their level of proficiency in both tests.

6.3.1 English Level Tests

The English levels, assessed with two different level tests from an online platform (Exam English[2]), were divided in Vocabulary/Grammar and Listening, and were part of phase I of this study (phases I and II will be presented in the next subsection). Each of them takes about 15 minutes to complete. The website is a free platform for people willing to study for an English exam. It contains free tests and simulations of consolidated standardized tests, such as the Test of English as a Foreign Language (TOEFL), the International English Language Testing System (IELTS), Cambridge English Exam, and so on.

Other than material for specific exams, the Exam English (henceforth EE) page offers two level tests: 1) Grammar/Vocabulary (G/V), and 2) Listening. Each test is comprised by 15 multiple choice questions and takes about 10-15 minutes to complete. The Listening test required the use of headphones, to increase the Sound-to-Noise Ratio (SNR). An example of how the tests looked like can be seen in Figure 10.

---

[2] The tests can be found at https://www.examenglish.com/leveltest/index.php (last access on April 20th, 2018)

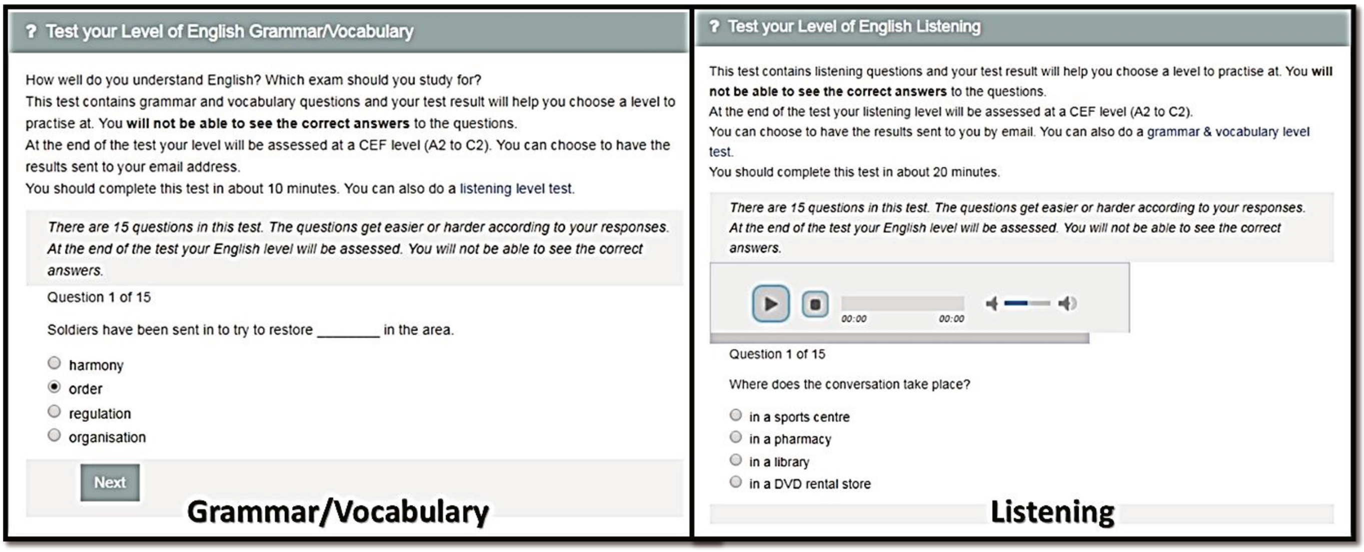


Figure 10: Grammar/Vocabulary and Listening tests.

After getting the official scores, the participants were supposed to send them directly to the researcher’s email. The English tests are assessed at a Common European Framework of Reference for Languages (CEF) level, a scale that provides a common basis for different languages syllabuses, curriculum guidelines, textbooks, etc. The assessment is supposed to happen in a comprehensive way, combining different skills judged proper to the effective use of language for communication, and, that way, overcoming barriers to communication among professionals of modern languages that may arise, given different educational systems (COUNCIL OF EUROPE, 2001).

There are six different levels of proficiency (A2, A1, B2, B1, C2, C1) that can be interpreted from the CEF scale, presented in Table 5. At the end of the tests, one of the CEF levels was calculated for all subjects.

Table 5: CEF levels and descriptions from the EE website.

C2	*The capacity to deal with material which is academic or cognitively demanding, and to use language to good effect at a level of performance which may in certain respects be more advanced than that of an average native speaker. Example: CAN scan texts for relevant information, and grasp main topic of text, reading almost as quickly as a native speaker.*
C1	*The ability to communicate with the emphasis on how well it is done, in terms of appropriacy, sensitivity and the capacity to deal with unfamiliar topics. Example: CAN deal with hostile questioning confidently. CAN get and hold onto his/her turn to speak.*
B2	*The capacity to achieve most goals and express oneself on a range of topics. Example: CAN show visitors around and give a detailed description of a place.*
B1	*The ability to express oneself in a limited way in familiar situations and to deal in a general way with nonroutine information. Example: CAN ask to open an account at a bank, provided that the procedure is straightforward.*
A2	*An ability to deal with simple, straightforward information and begin to express oneself in familiar contexts. Example: CAN take part in a routine conversation on simple predictable topics.*
A1	*A basic ability to communicate and exchange information in a simple way. Example: CAN ask simple questions about a menu and understand simple answers.*

The participants had to submit scores for *both* exams along with the responses of the questionnaire from phase I in order to be considered for this study.

6.3.2 Group Placement

Since we divided our participants in four different groups of proficiency, we had to normalize the six levels CEF scale for both G/V and Listening tests. The final groups (N1, N2, N3, N4) correspond to an ascending level scale (N1- least proficient, N4- most proficient) and is how we are going to refer to the groups from now on. As can be seen from Figure 11, there were more strict criteria to place someone in N3 and N4 (Listening G/V for N4 and G/V for N3), whereas N1 and N2 were more permissible for both G/V and Listening. Since this is a study within acoustic phonetics, it was decided to prioritize the Listening task over the G/V when placing subjects into the respective groups. Additionally, if the G/V and Listening scores differed significantly, therefore not fitting in the scale presented in Figure 11 (e.g. G/V= C1 and Listening= A2), the subject would not be recruited for phase II of this study.

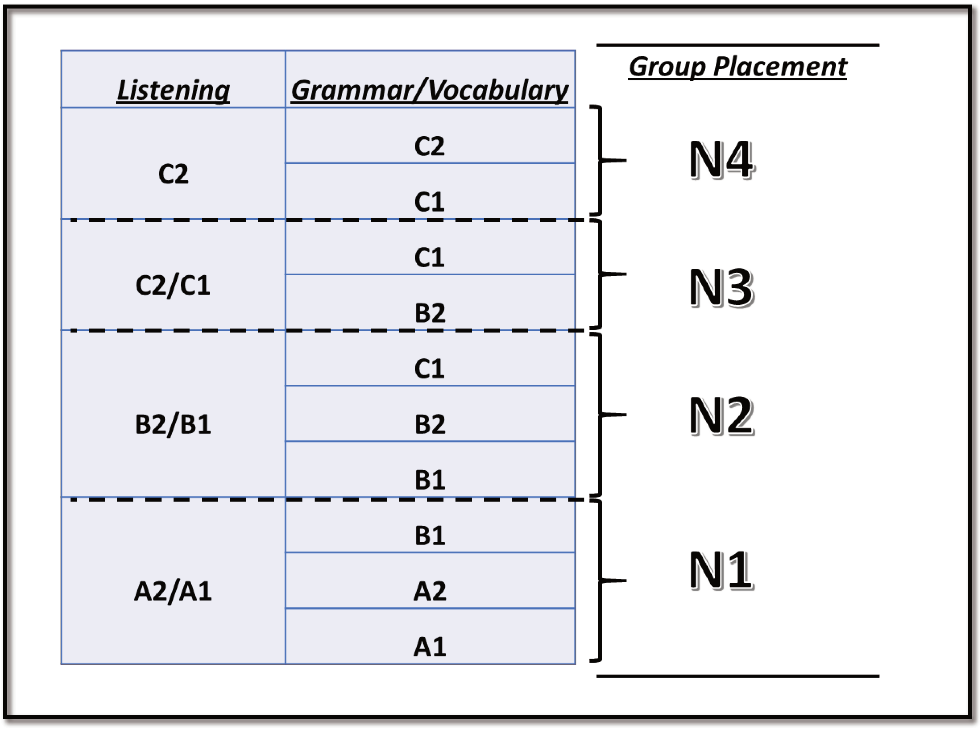


Figure 11: normalization of CEF levels for this study.

6.4 Screening Subjects

Initially, an online questionnaire developed on Google Documents (Figure 12a) was sent via email to a slew amount of student to select volunteers among those who met the inclusion criteria in a first basis. About 70 people responded to the questionnaire. The questions aimed to reveal a simple sociolinguistic profile of the participants. A printed version of the questionnaire can be found in the APPENDIX I. The responses were stored in a personal online database and an objective English level assessment test was sent to those who met the criteria. Along with the responses, the subjects also had to submit their scores from the EE tests, so that the first group placement could be done for each participant (if they fell under a possible group, as discussed in the Group Placement subsection).

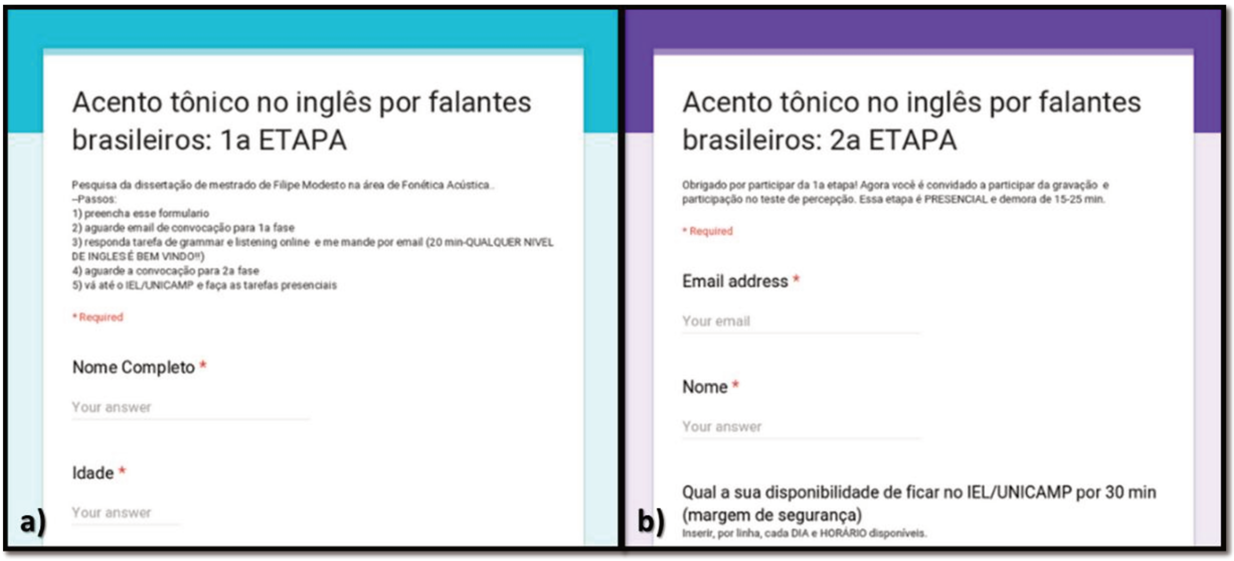


Figure 12: online questionnaires for phase I (a) and phase II (b).

Once the subject was placed in a certain group, the email recruiting him or her for the production and perception tests at the Phonetics Lab of Professor Dr. Plinio Barbosa (phase II- Figure 12b). After according to a time to meet at the lab, the final sample, comprised by 29 speakers, will be presented in the subtopic below.

6.5 Final Sample

All subjects whose data were used in this research met the inclusion criteria already mentioned in this section: being a post-lingual subject, with BP his/her first language, not had grown up in a bilingual environment, had submitted both G/V and Listening official EE scores, and being familiar with AE to at least some extent. The subjects who volunteered to participate either heard about the project via an online recruitment sent by university staff or directly contacted by the researcher. It is important to mention, that it was prioritized subjects out of the field of language studies, given our work was directed to knowing how the prosodic features happen in natural speech, by people who are not as aware of intrinsic linguistic processes as linguists are (KOCHANSKI, 2006).

Over 49 people, who fit in the inclusion criteria, responded to the EE tests and submitted their scores. Of these, only 29 were selected to participate in the study. The subjects were chosen based on their commitment to sending their score reports in time, and the scores themselves. They were carefully distributed into the subgroups, in an attempt to achieve a considerably homogeneous distribution in regards to number of subjects per level and sex within each level. It was also taken into consideration how convenient it was for the subject to go to the laboratory.

The subjects were distributed in four different groups based on their proficiency: N1, N2, N3, and N4. Information from the sociolinguistic questionnaire is shown in Table 6.

Table 6: subjects' description.

		Age	SEX		LIVED IN ENGLISH SPEAKING COUNTRY		YEARS OF STUDY
	Total (n)	*mean (years)*	*Female*	*Male*	*Yes*	*No*	*mean (years)*
N1	6	26	3	3	0	6	1.2
N2	8	21	4	4	1	7	2.7
N3	7	25	5	2	2	5	6.4
N4	8	26	4	4	2	6	7.4

As we can see from Table 6, the number was fairly homogeneous along the proficiency levels, with more females than males only in N3. The mean age of the participants was 24 years old (min=18; max=54; median=23). Only five of the 29 subjects had lived abroad in an English speaking country, for at least 6 months, and four of these five belonged to levels N3 and N4. We can see that the more proficient subjects were, as expected, the ones with more years of study. One of the questions in the sociolinguistic questionnaire asked the participant to give a score to his or her general level of English, in an ascending 1-4 scale so that we could compare to the actual level they were placed based on the EE exams (Table 7).

Table 7: tabled data of levels of English according to the self-assessment (S1-S4) and the measured EE tests (N1-N4).

	N1	N2	N3	N4
S1	1	0	0	0
S2	5	6	1	0
S3	0	2	4	2
S4	0	0	2	6

Most of the subjects considered themselves to be somewhere around the actual level of proficiency found through the objective exams. A few N3 subjects rated themselves as being N4 (n=2), and so did two N2 speakers of being N3. The majority of N1 speakers (five out of six) judged their levels better than found in the EE tests.

6.5.1 American Subjects

In total, 10 American subjects (7 female, 3 male) participated in this study. However, not all 10 participated in both tests (production and AR). One of them was recorded in Brazil, but had to leave before the AR test was finished, so we could not collect his data. Besides, in order to balance the amount of female/male ratio, so that we would not bias the results, only three subjects of each sex were used to analyze the acoustic parameters. The other nine subjects were recorded in Ithaca-NY. Their mean age was 24 years old, 50% was from the Northeast (New York), 20% from the Midwest (Illinois and Minnesota), and only one subject from the south (Georgia). None of them was balanced bilingual, but three had knowledge of other languages, including Korean, French, Spanish, Russian, and Portuguese.

The subjects were recruited directly by the researcher or via official email sent by the Department of Speech Language Pathology and Audiology of Ithaca College, once the study was approved by the institutional review board.

6.6 Experiments

There were three category of experiments carried out: 1) a production test (Brazilian and American subjects), 2) perception test (Brazilian subjects only), and 3) an accent rating test (American subjects only).

6.6.1 Production Test

The first step of phase II was the production test. The subjects presented themselves at the Phonetics Lab, and the following instructions were given: a slide presentation will be started with a sentence in each slide. You have to read the whole sentence in your habitual voice tone, speed, as if you were just ordinarily reading those sentences. There is no right or wrong. If you stumble on words, or hesitate or do not feel comfortable about the way you read it during a sentence, feel free to read it again.

The 73 words described in the Corpus subsection were presented with the assistance of a computer in a slide presentation within the carrier sentence: say______ again. The subject himself/herself would be responsible to scroll down to the next slide as soon as he/she finished the previous sentence. The microphone was positioned around 30 centimeters in front of the speaker, involved by a soft sheet in order to absorb reverberation and preserve the acoustic signal. The production test was recorded with the device: Zoom H4n Pro recorder. The recordings were made at a sampling rate of 44.1 kHz, and 16-bit quantization. A scheme representation of the production experiment is shown in Figure 13. In the picture, the slides are side by side to each other for better understanding of the task, but when the experiment was running, they were in full screen presentation mode, and randomized for every new subject.

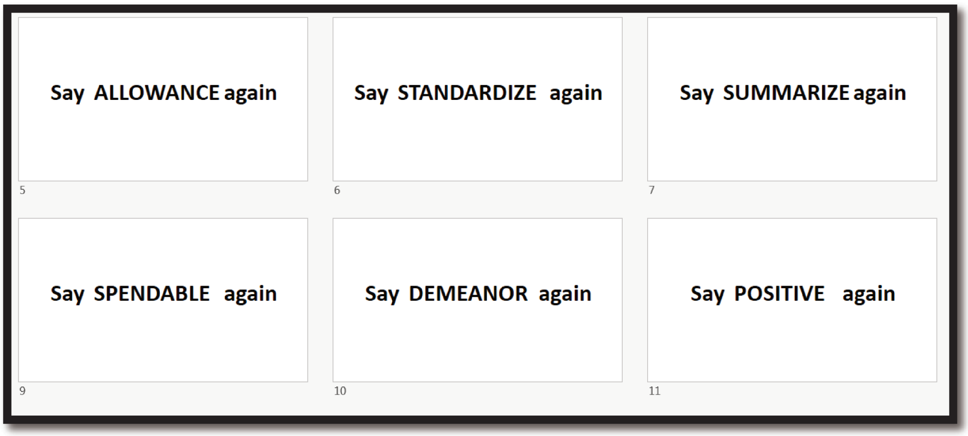


Figure 13: scheme representation of the production experiment.

All sentences were recorded continuously to avoid making the recording process more unnatural, and to optimize time. Every recording was stored separately for posterior acoustic analysis. The detailed analysis will be discussed in section 2.7.

The production test for the AE participants was recorded with the same device used with the BP speakers.

6.6.2 Perception Test

After the production test, every subject received instructions about the perception test. Still with the assistance of a computer, the Praat software was used to run the experiment, with a script developed by Barbosa (2017). Figure 14 gives a representation of the screen during the experiment.

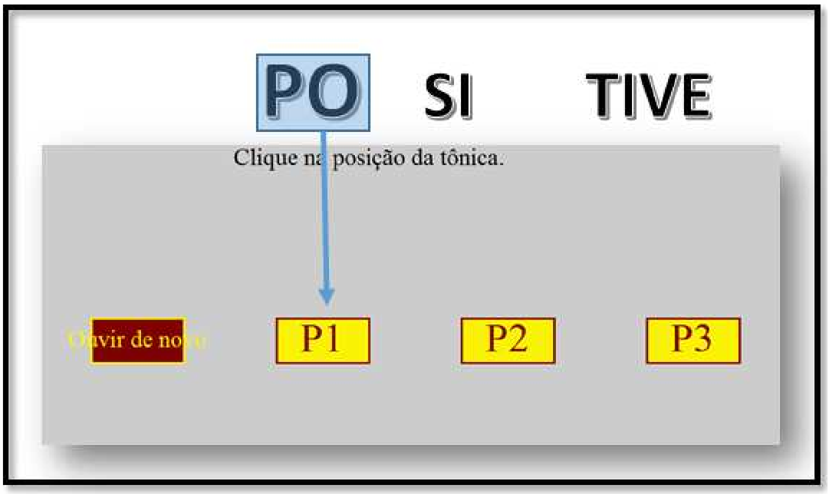


Figure 14: representation of the perception test.

Supra-aural headphones were placed on the subject’s head and the test would only start after a click on the screen. The instructions for this task: “you are going to listen to an amount of words (one at a time), and you have to click on the square that better match the position of the stressed/strongest syllable you perceive. You do not have to know what the word means and the words will repeat throughout the test”. A short practice was done before the actual experiment to make sure the task had been fully understood.

The same 73 words were presented out of the carrier sentence, spoken by a male speaker of 22 years old at the time, from Fairbury, Illinois. As soon as a square was clicked on, the next stimulus would start. Each stimulus could be heard again once (red button on Figure 14), but the test takers were encouraged to use that only if they were distracted for a moment or interrupted somehow. The experiment was paused after 60 stimuli and a short untimed break would start, for the subject would not get too tired from the task. All words were presented three times (total 219 stimuli) at random times so that we could test the consistency of the responses for each stimulus. If the same stress position were selected at least two times, that would be considered the final response. If three different positions were assigned for the same word, we invalidated the response for that stimulus. It is important to notice that Figure 14 shows positions P1, P2, and P3 as being the first, second, and third syllable position, respectively, because it is more intuitive for people to think of the positions according to the order of appearance. Since in phonetic studies, stress positions are often reported beginning from the end of the word, the test script was designed so that when someone clicked on the P1 button it would be printed as P3 in the final table for us, and the clicked P3 would be printed as P1.

At the end of the test, a ".txt" file was generated with the responses for all 219 stimuli. A script was developed using the R software to extract the mode of each stimulus and then write it in a new file.

6.6.3 Accent Ratings

Once all production data was properly recorded and analyzed, recording samples of BP speakers' of all four levels were collected to compile accent rating tests to be performed by the AE speakers of this study. A short practice was done before the actual experiment to make sure the task had been fully understood.

All AE speakers did the accent rating test after completing the production task. They were instructed to listen to the stimuli and rate each stimulus in a progressive 1-5 scale, 1 being the lowest and 5 being the highest score (Figure 15). All AE speakers had to take the same test twice. Each time a different aspect was being rated: Intelligibility and Nativeness. The former was described to the participants as being how intelligible the sample was, considering how easy it was to understand what was being said by each speaker. The latter should concern exclusively how the utterance sounded like produced by a native speaker of English. The test takers could replay the same stimulus once after each presentation, but they were encouraged to do so only if they were distracted for a moment or interrupted somehow. The experiment paused after 50 stimuli for a short untimed break.

Figure 15: scheme representation of the accent rating test for nativeness.

The test was comprised by 50 utterances: 10 from each level (total= 40) and 10 from an AE speaker (to see if the subjects were able to differentiate a native production among the BP speakers). The stimuli were randomized and played 3 times each to test the consistency of the responses, totalizing the 150 utterances of the task. If an utterance was rated equally at least twice, that responses would be recorded, whereas if three different rates were attributed to the same stimulus, the response would be invalidated. The utterances used for the task were the whole carrier sentence, since accent is easier perceived at an intonational level, not on each word individually.

At the end of both Nativeness and Intelligibility tests, a ".txt" file was generated with the responses for all 150 stimuli. A script was developed using the R software to extract the mode of each stimulus which was written in a new file.

6.7. Data Analyses

6.7.1 Acoustic Analysis (production test)

After recording all the subjects (Brazilian and American), each audio was segmented in four different tiers: Word, Syllable, Stress, and CorrectSP (Correct Stress Position), as depicted in Figure 16.

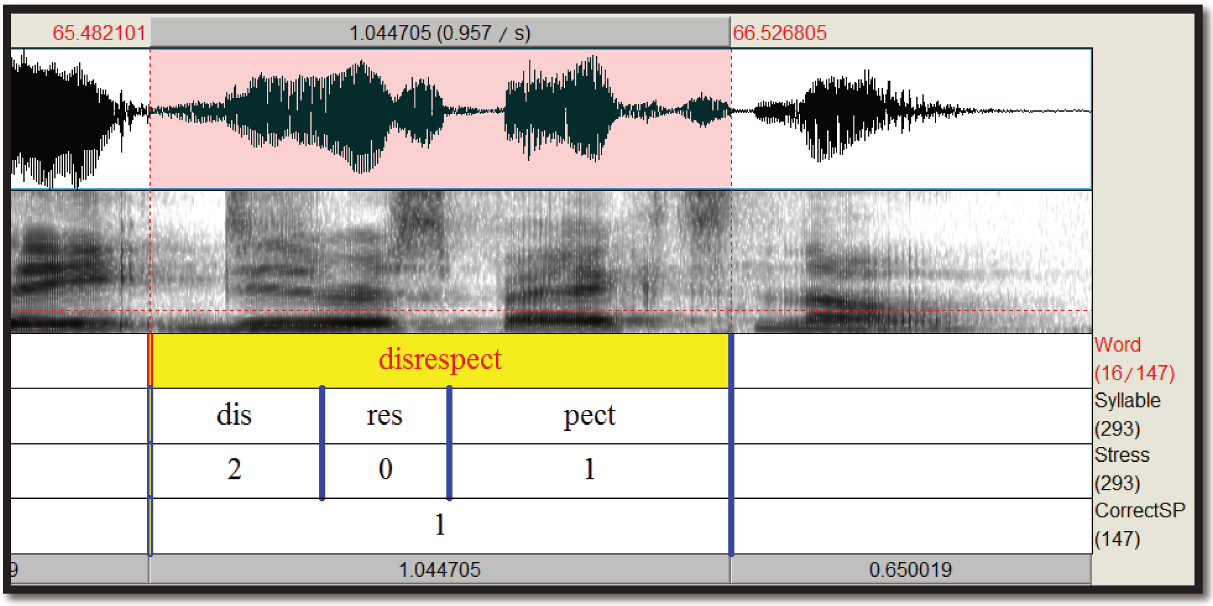


Figure 16: examples of tiers for each audio.

The acoustic analysis was carried out through the free software Praat (WEENINK, BOERSMA, 2018). The "Word" tier contained the word in focus (our target) of each carrier sentence. The "syllable" tier divided the word space in all the three syllables each word was comprised. The Stress tier had the exact marks of the "Syllable" tier but instead of the syllables, indexes were attributed to the kind of stress each syllable beared: 0= unstressed, 1= primary stress, 2= secondary stress. Finally, the 4th tier ("CorrectSP") indicated the correct stress position of the word, whether the participant succeeded in assigning stress correctly or not. It was used merely to help us compile the tables for the statistical analysis. In Figure 16, the "1" means it is a P1 word, a.k.a. bearing stress in the first position, from the end.

Once all audios were properly segmented, a script (developed by Barbosa, 2017) was run through them to extract the acoustic parameters of interest for this project for each segmented syllable. The parameters were: median of the fundamental frequency (F0MED- in Hz), peak of the fundamental frequency (F0PEAK- in Hz), standard deviation of the fundamental frequency (F0SD- in Hz), duration (DUR- in milliseconds), total intensity (TOTINT- in dB), and relative intensity (RELINT- in dB). This last one is the difference between the total energy up to the Nyquist frequency and the energy up to 500 Hz. It corresponds to the vocal effort employed when pronouncing a stressed syllable (SLUIJTER, et al. 1996). The appropriate frequency parameters were applied when extracting the acoustic parameters, given the physiological differences in the values of F0 for men and women. The F0 floor and ceiling used were 70-150 Hz for men and 120-500 Hz for women, respectively. Overall intensity is a parameter used to stress syllables in AE, when the word is in focus condition (such as in this study), therefore, it is of our interest to assess this parameter, and to analyze a possible difference between total and relative intensity in word stress in the English of BP speakers. F0 was carefully measured in different ways (peak, median, and standard deviation). It is known the relevance F0 has at the intonational level, but we were interested in seeing its role in linguistic stress.

6.7.2 Perception/ Accent Rating data

Both the perception and the accent-rating test were the same in nature when it comes to their analysis methodology. In both tests, the stimuli were presented three times at random during the test and each response was collected to a table, along with its response time. At the end of each test, the table was saved and stored. The free software R was used to read the tables and extract the mode of the three responses per stimulus (Figure 17). If there was no mode, that particular token was discarded.

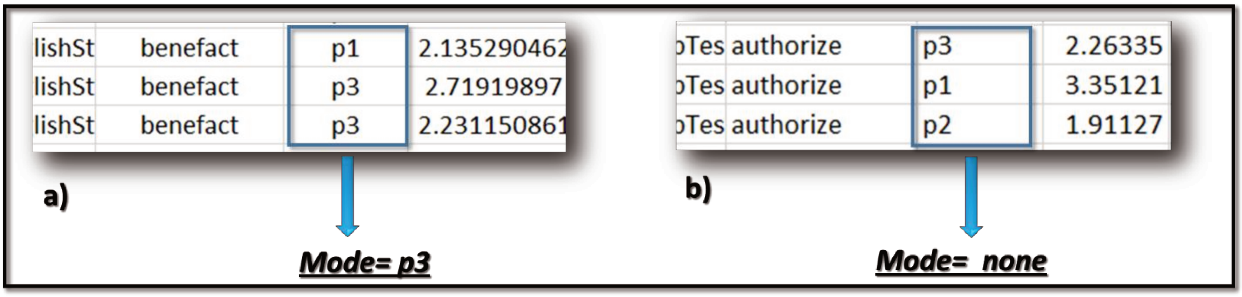


Figure 17: representation of how the responses in the perception and accent rating tests were collected.

In the example from Figure 17a, the stimulus *benefact* was perceived as having initial stress (P3 position), two out of three times for that participant. In Figure 17b, however, the stimulus *authorize* was perceived with a different primary stress in each presentation, indicating that the participant may have been distracted at the time or that he/she could not really point out what the perceived stressed syllable was. The response in the latter case was not computed.

6.7.3 Statistical Analyses

All data from this study underwent statistical analysis, either descriptive, inferentially, or yet both. Regarding the production task of both BP and AE speakers, raw acoustic data was analyzed in terms of mean, median, and standard deviation. The percentages of right answers in regards to stress position were carefully analyzed according with proficiency level, cognate relationship, and preferred stress position (if any). The inferential analyses were carried out with the raw data, but also with the normalized data of all acoustic parameters.

Acoustic data normalization is necessary whenever comparing multiple speakers. The reason for that is that there is a lot of variation in speech that is related to intrinsic aspects of the speaker, and depends on articulatory (anatomical/physiological) variation. Therefore, in order to obtain more correctly classified values for multiple vowel tokens (vowel- extrinsic) Lobanov's z score transformation (1971) was applied. Lobanov's procedure is considered one of the most successful in a perceptual perspective (ADANK et al, 2004).

In order to identify which acoustic parameters were used to realize linguistic stress, several 2-way ANOVA models were performed. Since the data failed to meet the three conditions for the conventional ANOVA (normality of the residuals, homoscedasticity of variances, and independence of the samples), a non-parametric method had to be used instead, called Scheirer-Ray-Hare (SHR). The test is an extension of the Kruskal-Wallis test, and a significance level of 5% was considered throughout the tests.

The analyses took into consideration the English proficiency level (PL), cognate status of words, correlation between production and perception, correlation between perceived accent by the AE speakers and the PL and the acoustic parameters used to stress syllables.

Duncan's multiple range test used the minimal significant difference to analyze whether the acoustic features of BP speakers' production could be grouped independently from one another, and which level would be closer to the AE speakers.

The *Effect Size* (ES) was calculated to understand how much each of the independent variables (stress level and cognate status) contributed (effected) to explain the variance of the dependent variables (the acoustic parameters) per PL.

6.8 Ethics Aspects

The steps of this study involved participants living in Campinas-SP (Brazil) and Ithaca-NY (USA). Before beginning data collection this project was submitted and approved by the institutional review board of each institution: Ethics Committee of the University of Campinas (Comitê de Ética em Pesquisa- CEP/UNICAMP), under the protocol CAAE: 66786617.2.0000.5404, and the Institutional Review Board for Human Subjects Research of Ithaca College, under the protocol IRB 0118-09, number 00004870.

7 Results

We obtained results for the three major tasks from this study: production, perception, and accent rating tests. Each of these will be discussed in further details in the next subsections.

7.1 Production

The data from the production test (recording the participants reading the list of words within the carrier sentence) was analyzed acoustically, so that we could understand how lexical stress is realized by the BP and AE speakers, therefore allowing us to make comparisons between both populations (and BP subgroups N1-4). At first, the acoustic data are presented, followed by the percentage of scores (hits) for each level overall, specifying the cognate status (CS), and the native stress position (SP) of the words. Finally, the statistics involving the acoustic data is performed so that we could investigate which acoustic parameters were used to stress syllables in each group, how they differ among each group, and what is the contribution of each parameter to syllable stress.

7.1.1 Acoustic Data

The recordings from the production task were segmented and analyzed in Praat. A script was used to automatically extract 12 acoustic parameters from the segmented intervals from each tier. It is well known that duration, fundamental frequency, and intensity are the three main parameters used to describe prosody. For that reason, we based our analyses on six of these 12 parameters: duration (DUR), median fundamental frequency (F0MED), standard deviation of the fundamental frequency (F0SD), peak of fundamental frequency (F0PEAK), total intensity (TOTINT), and relative intensity (RELINT). It is important to understand what the indexes related to stress mean: 0= unstressed syllable, 1= syllable with primary stress, and 2= syllable with secondary stress. The five groups represent the four levels of English proficiency (N1-4), plus the American control group (USA).

The average values for each parameter are presented below in the series of tables (Table 8 through 13). In addition, the three right-most columns of these tables are displaying the differences between the means of each parameter across stress conditions (0, 1 or 2). This allows us to explore not only which stress condition has more influence on the acoustic realization of stress, but also how the three stress categories (unstressed, primarily stressed, secondarily stressed) are related to each other.

Table 8: mean values for duration for the three stress levels (0, 1, and 2) for the five groups (N1, N2, N3, N4, and USA), and the differences among groups.

	DUR(ms)					
	0	**1**	**2**	**dif 1-0**	**dif 2-0**	**dif 1-2**
N1	211	338	247	127	36	91
N2	209	330	230	121	21	100
N3	209	311	266	102	57	45
N4	189	303	254	114	65	49
USA	169	243	213	74	44	30

Table 9: mean values for the median of the fundamental frequency for the three stress conditions (0, 1, and 2) for the five groups (N1, N2, N3, N4, and USA), and the differences among groups.

	F0MED (Hz)					
	0	**1**	**2**	**dif 1-0**	**dif 2-0**	**dif 1-2**
N1	179	176	196	-3.6	17.2	-20.8
N2	188	180	190	-8.6	2	-10.6
N3	207	197	195	-10	-11.3	1.3
N4	165	159	169	-5.6	4.8	-10.4

USA	154	158	145	3.4	-9.2	12.6

Table 10: mean values for the peak of the fundamental frequency for the three stress conditions (0, 1, and 2) for the five groups (N1, N2, N3, N4, and USA), and the differences among groups.

	F0PEAK (Hz)					
	0	**1**	**2**	**dif 1-0**	**dif 2-0**	**dif 1-2**
N1	194	196	219	1.9	25.1	-23.2
N2	204	201	208	-3.4	3.2	-6.6
N3	228	225	219	-3.8	-9.7	5.9
N4	182	177	185	-4.7	2.9	-7.6
USA	167	171	162	4.1	-4.9	9

Table 11: mean values for the standard deviation of the fundamental frequency for the three stress conditions (0, 1, and 2) for the five groups (N1, N2, N3, N4, and USA), and the differences among groups.

	F0SD (Hz)					
	0	**1**	**2**	**dif 1-0**	**dif 2-0**	**dif 1-2**
N1	8	10	12	1.8	4.3	-2.5
N2	9	11	10	1.6	1.2	0.4
N3	13	14	12	1.4	-0.3	1.7
N4	10	9	10	-0.4	0	-0.4
USA	8	8	9	0	1.4	-1.4

Table 12: mean values for the relative intensity for the three stress conditions (0, 1, and 2) for the five groups (N1, N2, N3, N4, and USA), and the differences among groups.

	RELINT (dB)					
	0	**1**	**2**	**dif 1-0**	**dif 2-0**	**dif 1-2**
N1	4	5	5	0.7	0.9	-0.2
N2	4	5	4	1.1	-0.1	1.2

N3	4	5	5	0.8	0.5	0.3
N4	3	4	4	1.1	0.2	0.9
USA	4	5	4	1.5	0.5	1

Table 13: mean values for the total intensity for the three stress conditions (0, 1, and 2) for the five groups (N1, N2, N3, N4, and USA), and the differences among groups.

	TOTINT (dB)					
	0	**1**	**2**	**dif 1-0**	**dif 2-0**	**dif 1-2**
N1	64	63	65	-0.8	1.4	-2.2
N2	65	64	65	-0.6	0.3	-0.9
N3	64	64	64	-0.5	-0.5	0
N4	65	65	65	-0.1	0	-0.1
USA	59	60	59	0.2	-0.3	0.5

The relative aspect of the means difference is better visualized with the help of the plots, in Figure 18. They represent the plots of the three right-most columns of the last six tables above presented. By looking at the duration plot in Figure 18, we see an overall decrease in the use of duration to assign stress, since all three categories (dif 1-0, dif 2-0, dif 1-2) tend to approximate to one another and also, the difference between primarily stressed and unstressed syllables is always the biggest in all groups. The order of priority (in duration) in the differences for the dif 2-0 and the dif 1-2 groups is the same between N1-2, and the same for N3-4 plus USA. Another plot in the same figure that showed an interesting fact is total intensity. The three plots for N1 are sparse from one another, and tend to funnel the sparsity along with level increase. The three differences analyzed for the raw F0 measurements (median, peak, and standard deviation) and RELINT measurements did not show a linear behavior throughout the levels.

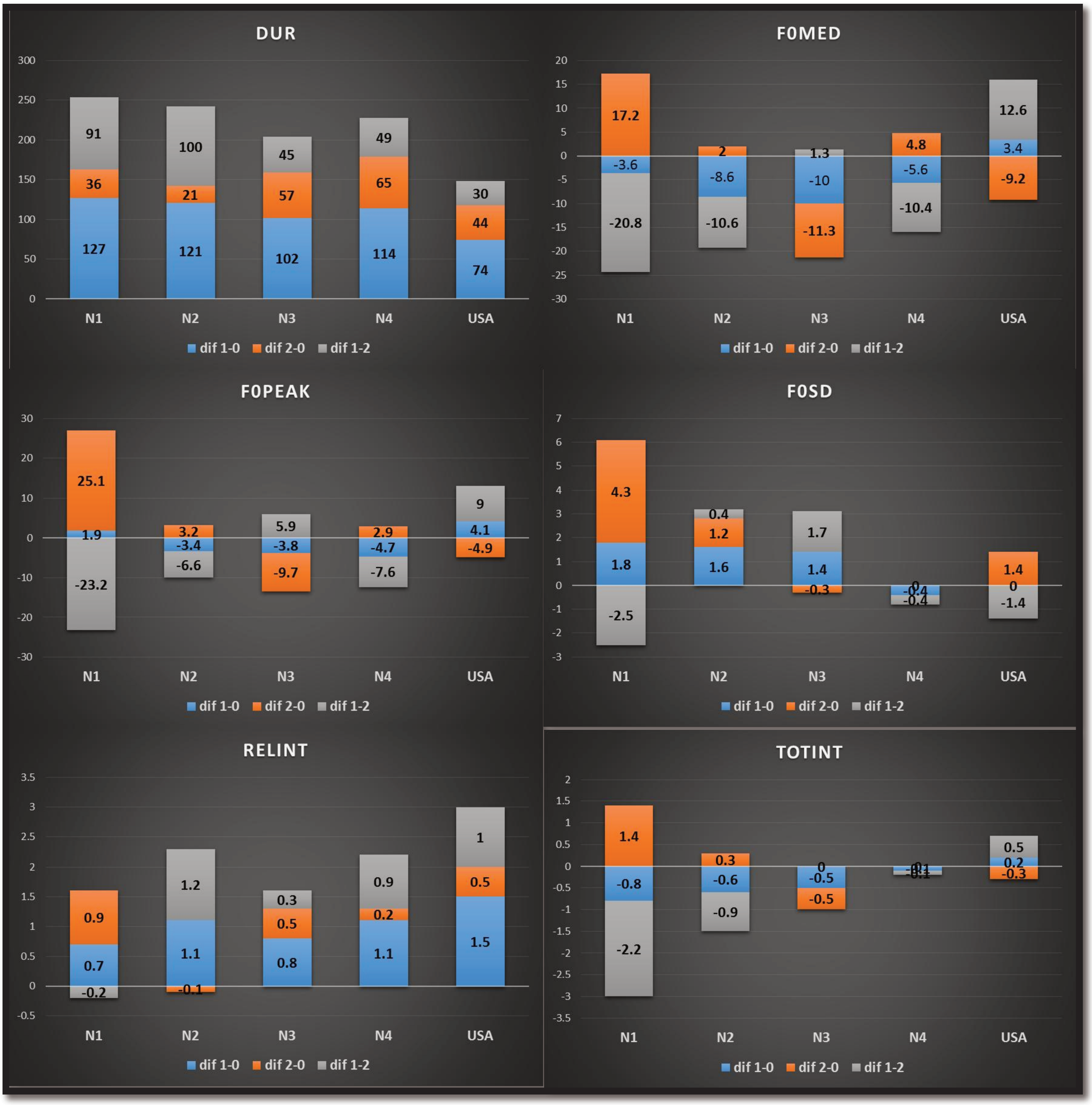


Figure 18: differences of means of stress condition: dif 1-0 (difference between primary stress and unstressed), dif 2-0 (difference between secondary stress and unstressed), and dif 1-2 (difference between primary and secondary stress).

Considering the analyses done so far were performed on raw acoustic data, it is important to understand the role that individual variances caused by anatomy and gender differences play in these results. For that matter, the procedure of data normalization (Lobanov's method specifically) can be very helpful eliminating individual and gender related influences without interfering in the phonemic properties of the material (ADANK, 2004). Lobanov's method (equation 1) puts each value in its relational position according to the average, and standard deviation of the productions for each subject s. In the equation, N_i is the normalized value, x_i is the raw value, μ_s is the mean, and δ_s is the standard deviation, computed for all parameter values for each subject.

$$Ni = \frac{xi - \mu s}{\delta s} \qquad (equation\ 1)$$

The mean values for these normalized measures across the three stress conditions are shown in Figure 19. The means are displayed comparing the entire Brazilian sample with the American sample for a major contrast between both populations. Then, we calculated the means for each individual level, to analyze their contribution separately. That way (with the normalized data), we can compare the means and their differences among groups without having to worry about the constraints previously mentioned.

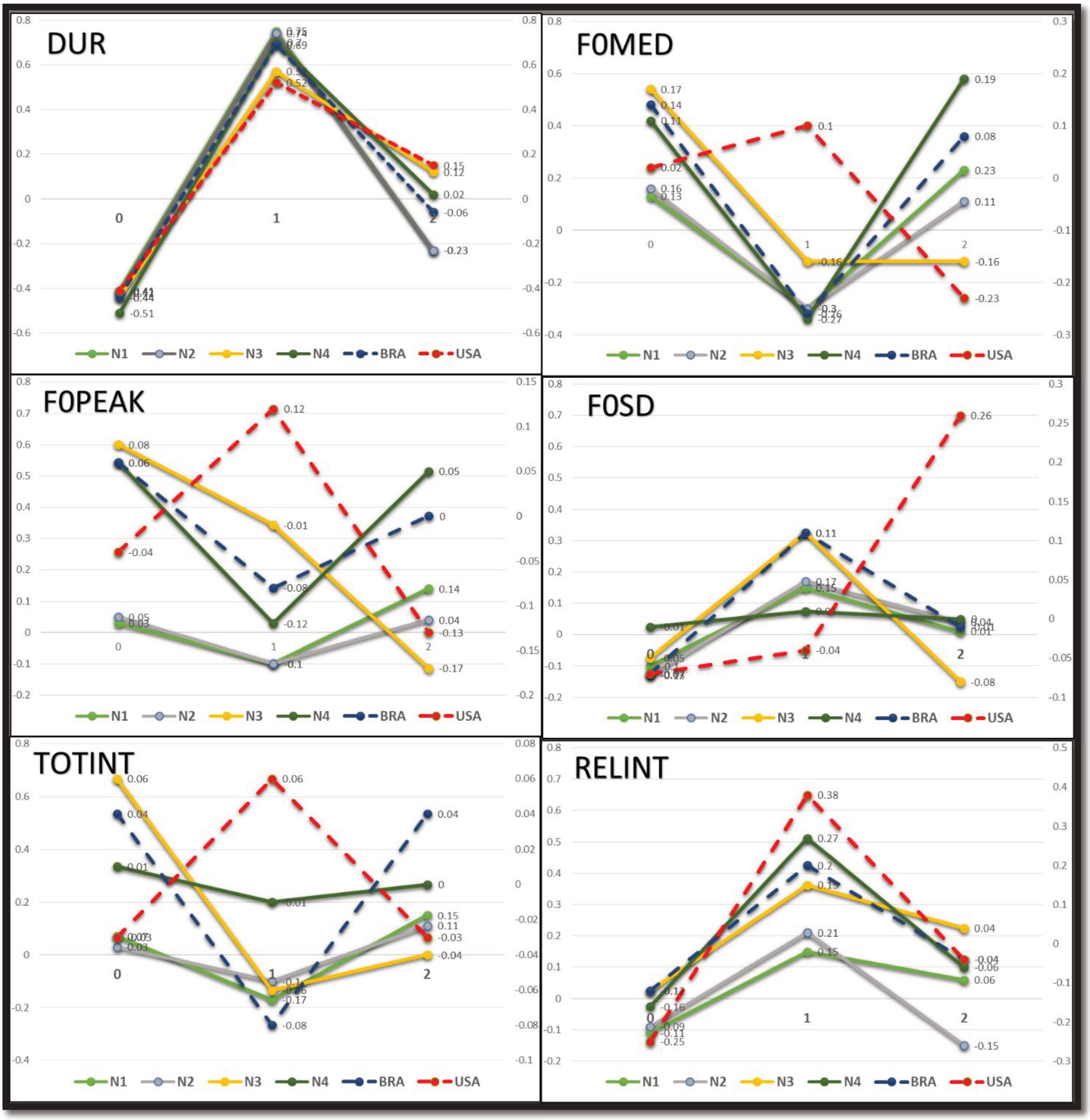


Figure 19: plots of normalized means of acoustic data per stress position (0, 1, 2) per level (solid lines) and the entire BP population (dashed blue line) and AE population (dashed red line).

Duration seems to be the most notorious plot in regards to a certain parallelism along levels N1-4 and consequently, a proximity with the AE population. The N1-2 groups present greater distance between means of duration of syllable with primary and secondary stress. Speakers from N3-4 and the AE population decrease the difference between the two stress categories for the parameter. Means for the unstressed syllables remained concentrated in about the same area for all five groups. Another parameter that follows the pattern for duration was relative intensity (RELINT). The parameter has its highest mean in primarily stressed syllables for all groups, but the mean increases along with level, and the AE speakers present the greatest values for the parameter.

Despite some particularities, TOTINT and the the F0 measures (except F0SD) seemed to behave the same way: the plot for stress positions 0, 1, and 2 for the AE population had the greatest mean values for 1 in comparison with 0, and 2, whilst the four BP groups behaved the opposite way, with the lowest values in 1. The plots of both populations in these parameters look like two inverse triangles overlapped. When considering such inverse patterns in these parameters, no BP groups approached the native AE speakers.

7.1.2 Percentage of hits

Regardless of the acoustic parameters used to stress syllables, we analyzed which syllables the participants made prominent in their production. That allowed us to calculate the amount of correctly stressed syllables each group had. We cross-referenced the position of the syllable stressed by the participant within a word with the position where stress should have fallen. If the positions coincided that meant the participant stressed the syllable correctly. That is called a *hit* in this study. Figure 20 shows the overall percentage of hits for the four levels.

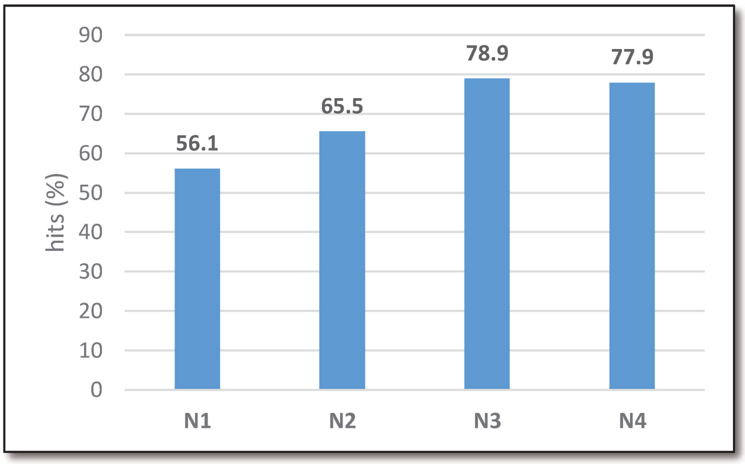


Figure 20: overall hits for N1-4 (production).

The overall results show that there is a tendency of increasing scores along with the proficiency level, which makes sense in a preliminary widespread analysis: the higher the proficiency level in the language, the higher the rates of correct answer (even though N4 is slightly below N3). Below (Figure 21), the hits are shown divided in C (cognates) and NC (non-cognates), per level.

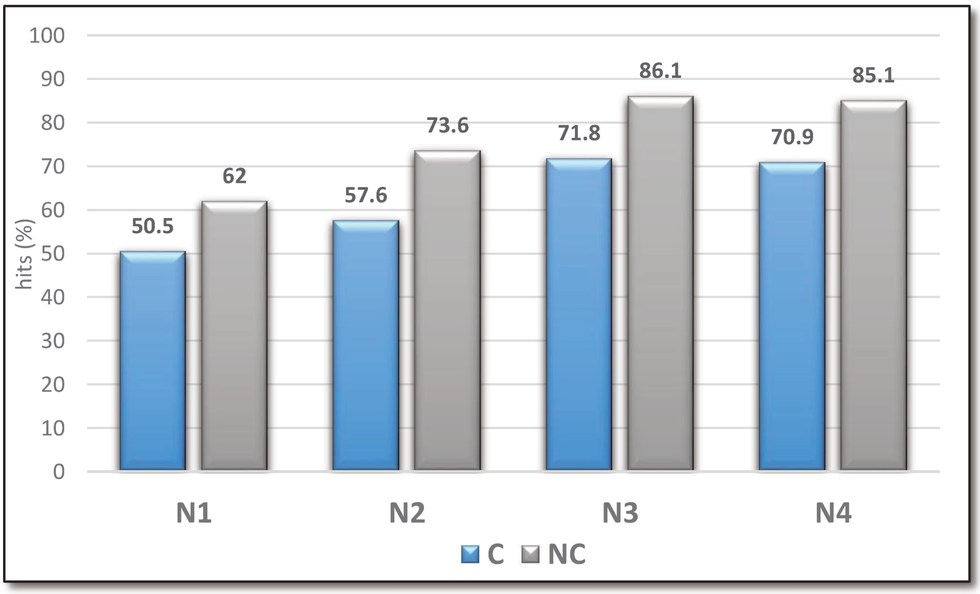


Figure 21: hits per cognates per level (production).

For all four levels, the non-cognate words were easier to get the right stress placement than the cognate ones, even though such difference (hits COG – hits NON COG) seems to be about the same through all four levels: dif N1= 12, dif N2= 16, dif N3= 14, dif N4= 14. The overall percentage, with all four groups combined, was 63.2% for cognates and 77.3% for non-cognates. This higher percentage of NC hits corroborates one of our hypothesis that the production of cognate words is much more prone to be triggered by the L1 correspondent than a non-cognate, leading to a major mispronunciation of the cognate words in comparison to non-cognates.

In order to investigate if there was an easier stress position for the Brazilian sample, we analyzed the hits according to each of the three word positions that could bear stress (P3, P2, P1). The plots for that are in Figure 22.

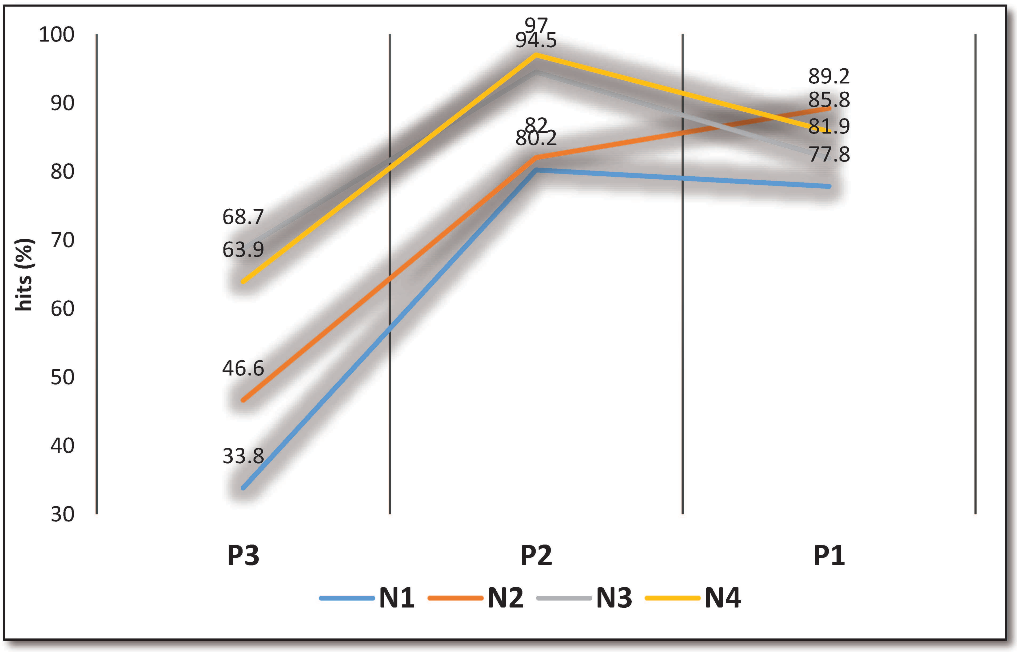


Figure 22: hits per position per level (production).

Overall, speakers of all levels stressed correctly words where it is expected final stress. In this graph, we used all the words together. In order to analyze effectively the influence of L1 in L2 prosody we should consider only the cognate words, which are the ones capable of triggering an already known stress pattern for the BP speakers. This analysis is shown in Figure 23.

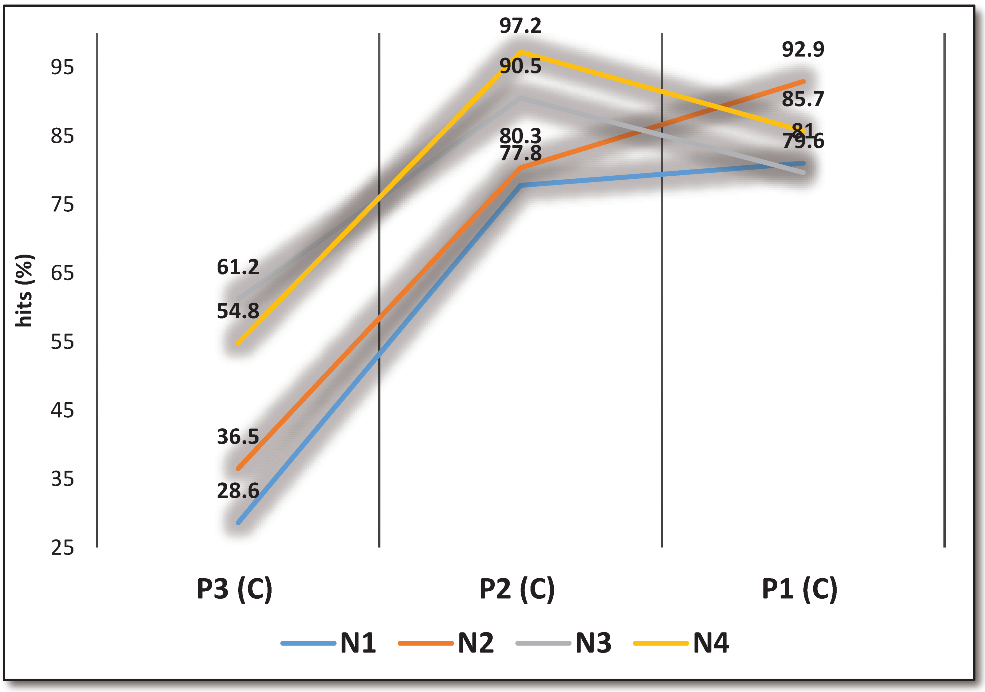


Figure 23: hits per position per level (cognates only).

If we analyze Figure 23, we might wonder why the N3 and N4 speakers seemed to decrease their progress from P2 to P1, while the N1 and N2 speakers seemed to keep increasing their scores. To look a little closer at that finding, we plotted the differences of hits between C-NC for each level to each position (Figure 24).

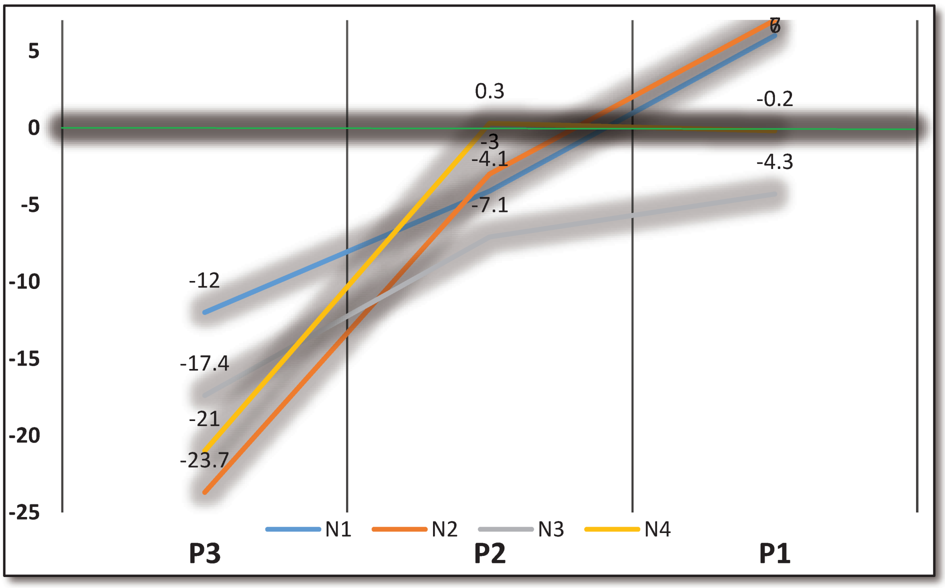


Figure 24: difference C-NC (in %) per level per position.

The C-NC differences shown in Figure 24 begin negative for P3, as the NC scores were higher than the C ones, expressing the effect of L1 in stress misplacement in that position. The cognates effects for N4 is reduced to almost zero for P2 and P1, which might be because the higher level speakers tend to incorporate the new L2 words to their phonological systems in a separate category, even the ones that are morphologically alike with their L1 (C-NC=0%). Even though that only happened for the first two stress positions, we could hypothesize that with more training the C-NC differences for P3 might also tend to zero, maybe placing each pair of C/NC together and making evident that they differ in stress (e.g.: *DEmonstrate* and *demonsTRAR*). N1-2 kept increasing their C-NC difference, now to the positive dimension in P1, now demonstrating a hypercorrection for the words where lexical stress is the same for both languages (e.g.: *interACT* and *interAGIR*).

In order to investigate how the cognates behaved differently according to stress position in each group, we broke down the hits per level and per cognate status. Figure 25 shows the hits for P3, Figure 26 for P2, and Figure 27 for P1.

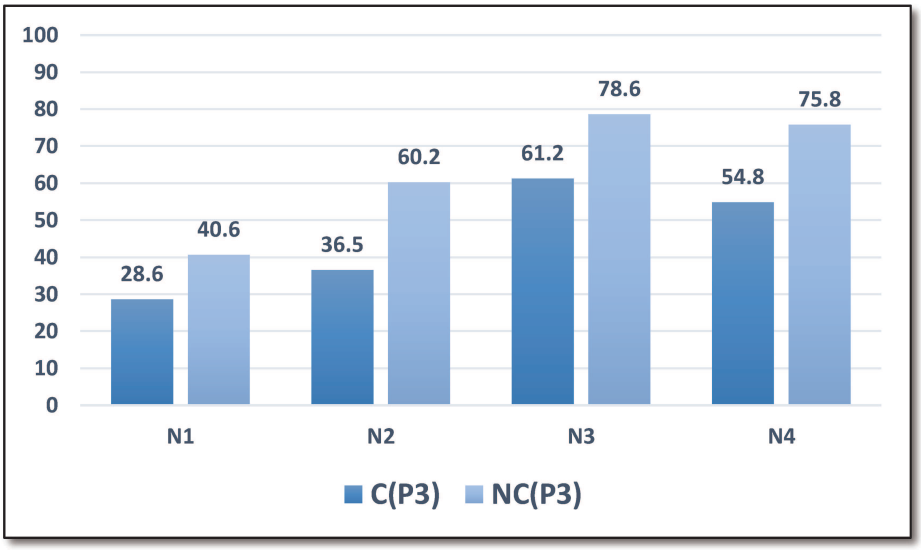


Figure 25: hits (in %) per cognate type per level for P3.

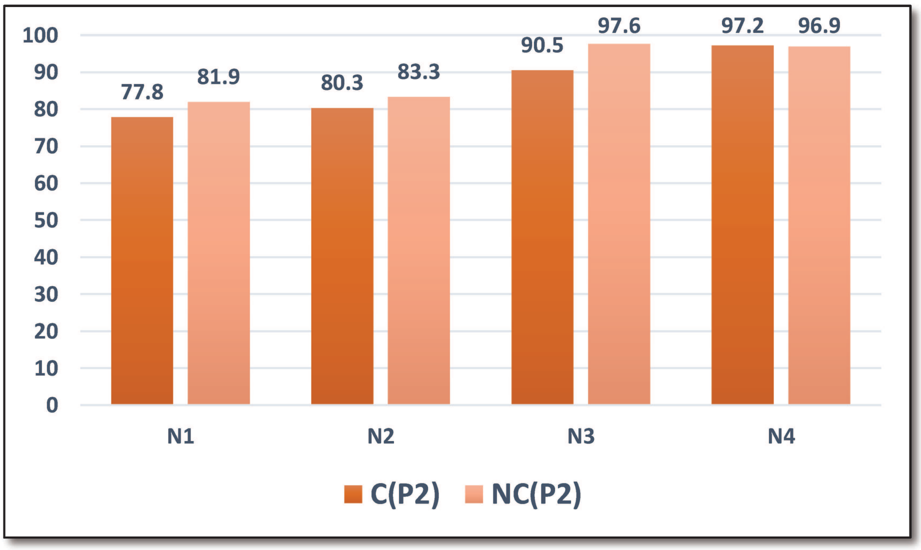


Figure 26: hits (in %) per cognate type per level for P2.

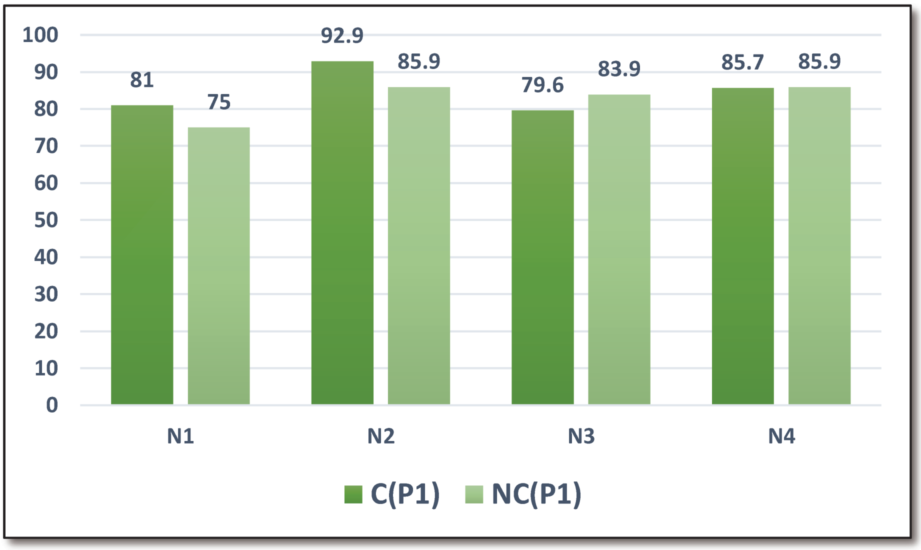


Figure 27: hits (in %) per cognate type per level for P1.

7.1.3 Analyses of Variance (ANOVA)

A series of two-way analyses of variance was used to study how the acoustic parameters in the production were affected by the factors stress and cognate status. Since the data did not meet the criteria for the parametric 2-Way ANOVA we had to use the SHR method, as described in the methods section.

The *STRESS* factor had two levels: stressed and unstressed. The former included syllables bearing both primary and secondary stress. The *cognates* factor also had two levels: cognates and non-cognates.

In Table 14, the results for the SHR model are shown. Significant differences between conditions would be indicated by an alpha below 5%. For all five groups, duration was used to assign stress, as well as relative intensity, correlate of vocal effort. The means of F0MED were also significantly different between stressed and unstressed syllable, with the lowest frequency values happening in unstressed syllables. Regarding cognate status, F0PEAK and TOTINT were affected by it. The highest values of F0 peak happening in the cognate words (in all five groups) and the highest values of total intensity happening in the non-cognate ones.

Table 14: results for the SHR model (AP~STRESS+COGNATE) of the normalized data. T= stress, C= cognates.

	DUR	F0MED	F0PEAK	F0SD	RELINT	TOTINT
N1	T	T	C	T/C	T	T/C
N2	T	T	T/C	T/C	T	C
N3	T	T	T/C	C	T	T/C
N4	T	T/C	T/C	C	T	C
USA	T	-	C	C	T	C

The result that cognate status posed a difference in the means of F0PEAK, F0SD, and TOTINT for the AE population may be in fact a result of word familiarity of the AE speakers to the words in the study. When establishing the corpus for this study, some difficulties emerged from the process of finding words that could be perfectly distributed with the desired phonetic properties. For instance, words that were non-cognates with BP, ending in –ate. This suffix is strictly related to words with a Latin root, therefore, also being related to BP. The ones that were not cognates had very low frequency of use in English (e.g.: abrogate, tessellate, adumbrate). That happens because most of the P3 words which suffix we tried to control for are in their nature cognates with BP and other Latin related languages. Forcing words that are non-cognates between BP and AE but still share the same morphological feature generates words in English that are unfamiliar even for the American participants, and this unfamiliarity may have been expressed as a non-cognate category within their own language, as the analysis of variance reveal.

In order to test this, the analyses were redone, selecting stimuli which were supposedly familiar to the native AE speakers. Such familiarity was assumed using the 20,000 more frequent words from the Google's Trillion Word Corpus (GTWC)[3].

Out of the 73 words from our study, 30 words were eliminated, 8 being cognates and 22 noncognates. The new SHR model for the remaining words for the AE population is described in Table 15, below.

Table 15: results for the SHR model (AP~STRESS+COGNATE) of the normalized data analyzing different word groups. T= stress, C= cognates.

	DUR	F0MED	F0PEAK	F0SD	RELINT	TOTINT
USA 1	T	-	C	C	T	C
USA 2	T	-	-	-	-	-

USA 1 presents the acoustic parameters with statistical significance using the entire corpus, retrieved from Table 13. USA 2 presents the new analyses, without the 30 infrequent words mentioned before. We can see that, as expected, the cognate status is no longer considered when stressing words for the AE speakers, and duration remains as a significant factor in both analyses.

[3] *The GTWC is a corpus resulting from an extremely vast analysis using n-gram models on Google datacenters to describe the trillion most used words in webpages, originally described by Franz and Brants in 2006. The corpus has been used in many statistical modeling applications, such as spelling correction, speech recognition, etc.*

7.1.4 Effect Size

After doing the SHR models, the effect size (eta square method- equation 2) of the acoustic data was calculated for each significant parameter in the statistical models, so that we could analyze how much of the variance of the parameter was explained by the factor (stress or cognate status). SS_{factor} is the sum of squares of the investigated factor, and SS_{total} is the total sums of squares.

$$\eta^2 = \frac{SS_{fator}}{SS_{total}} \qquad (equation\ 2)$$

Table 16 shows the results of the effect size calculated with the eta-square method. Clearly, stress has exerted more effect in duration than did any other parameter (in all five groups), explaining most of the variance for the parameter.

Table 16: effect size results for N1-4 and the AE sample for the six parameters studied.

	DUR	F0MED	F0PEAK	F0SD	RELINT	TOTINT
N1	19	3	-	1	1	1
N2	18	4	0	1	1	-
N3	17	4	0	-	1	0
N4	23	2	0	-	2	-
USA	16	0	-	-	6	-

7.1.5 Duncan's Multiple Range Test (DMRT)

The DMRT is a statistical analysis that uses a studentized range that compares groups of means (in our case, the acoustic parameters of each group) and classifies them together in subgroups according to the proximity of the results (DUNCAN, 1955).

Because we had to use the raw acoustic data for such analysis, the groups were divided according to gender to attenuate some of the acoustic differences inherent to physiological features. Table 17 shows the test results for females and Table 18 for males. The sequence of the letters in the test has a linear characteristic with their values, so if there is a group *d,* it is closer (in means) to *e* than group *a.*

Table 17: DMRT results for females.

	DUR	F0MED	F0PEAK	F0SD	RELINT	TOTINT
N1	a	a	a	b	a	a
N2	b	b	b	b	a	a
N3	b	ab	a	a	b	c
N4	b	c	c	b	c	b
USA	c	d	d	c	c	d

Table 18: DMRT results for males.

	DUR	F0MED	F0PEAK	F0SD	RELINT	TOTINT
N1	ab	a	a	c	d	b
N2	a	b	b	a	c	c
N3	ab	c	c	c	a	b
N4	b	d	d	b	c	a
USA	c	e	e	c	b	d

With the raw data of the female population, it is much more possible to establish a relationship of linearity than for the male population. For all parameters (except F0SD), N1 has the farthest distance from the native female American speakers, and the order tends to be one that level reaches out to the native production linearly. For the relative intensity means, in fact, the difference between N4 and USA was almost none, that is why the test placed them in the same group.

The order of the male groups, on the other hand, does not follow a particular pattern. Only for F0MED and F0PEAK the four BP groups and the USA group follow a linear order.

7.2 Perception

The most frequent results for each stimulus of the perception test were tabled and then the hits were analyzed overall, per level, considering the cognate status, and syllable position, as was done for the production test.

7.2.1 Percentage of hits

We cross-referenced the selected syllable position chosen by the participant with the position where the primary stress actually was. If the positions coincided that meant the participant had perceived primary stress correctly. That is called a *hit*, just like in the production test. The stimuli for this task were the same words used in the production test, read by a native AE speaker, but were always presented after the production test, to avoid phonetic cues interference in the subjects' pronunciation.

The overall hits (Figure 28) for the perception task show an increase in hits along with level, just like in the production test. The overall numbers are higher in this test (except N3 whose hits were 79% in the production test), so lexical stress is easier to be perceived than it is to be produced correctly. Despite differences in its acoustic realization, stress is a feature present in both AE and BP, making speakers of both languages sensitive to it, which explains the better results for perception than for production. An interesting analysis would regard the mistaken responses for each participant, that is, those who clicked in three different stress positions for the same stimulus that is, a lack of consistency in subjects' responses), in order to observe some sort of stress deafness. However, when computing all the wrong answers from the 29 subjects, we could only find 97 inconsistent responses amongst the 2117 sets of stimuli presented. This comprises only 4.5% of the presentations, making any inferences from this data biased. These 97 inconsistent responses were not used to calculate the percentage of hits.

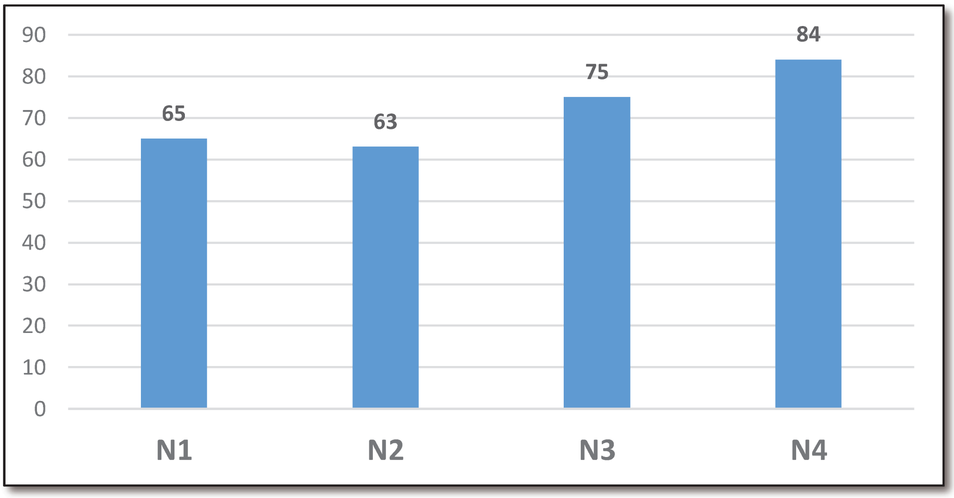


Figure 28: overall hits for N1-4 (perception).

When comparing the percentage of hits (Figure 29) we observe a much smaller difference between C-NC in all levels in comparison to the production test results. The perception of stress is related to the prosodic information in a speech signal. Since both BP and AE are stress languages with a non-fixed pattern, it is understandable that that kind of information is easier to be detected (FLEGE, 1995).

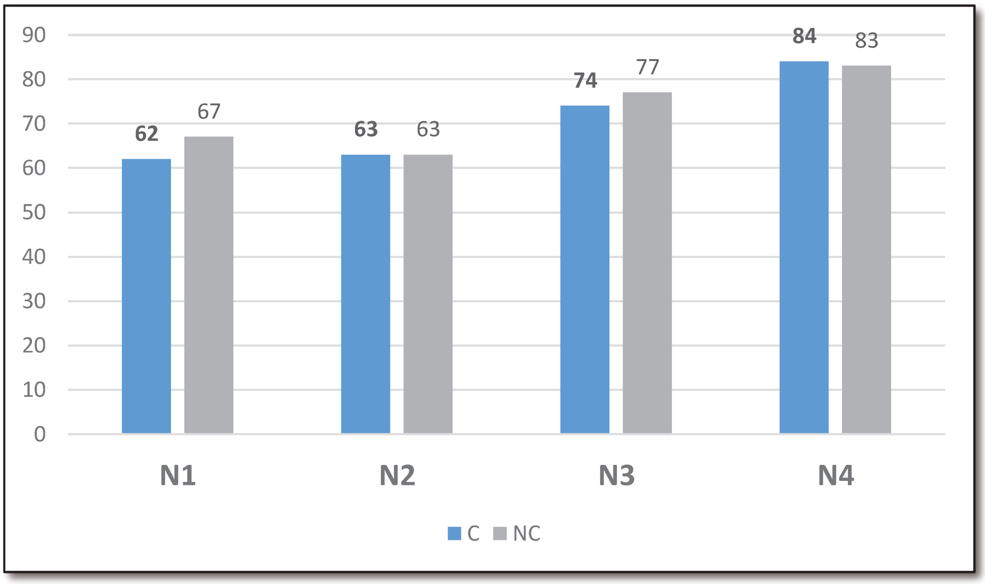


Figure 29: hits per cognates per level (perception).

The analysis per stress position (Figure 30) show that upper level speakers (N3-4) scored better than lower level speakers (N1-2) in all three positions, and that P2 was the easiest position to perceive stress of all. It makes sense since a P2 word is surrounded by a pre and post tonic syllables that are both unstressed. This difference between stressed (1 and 2) and unstressed syllables is well established. When secondary stress is present, though, the scores for all levels lower, especially for word with final stress. Words with initial or final stress (P3 and P1, respectively) both bear secondary stress, with alternating positions, that is, when primary stress is in the first syllable, secondary stress is in the last one (e.g.: suffocate), and vice versa (e.g.: interact). The perception of prominence may have been enough for the participants to click on a certain syllable.

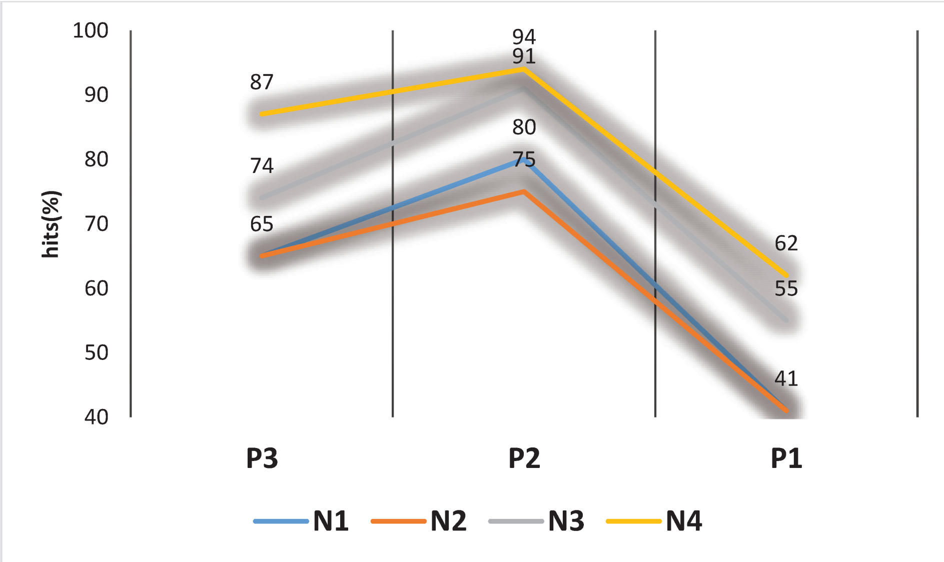


Figure 30: hits (%) per position per level (perception).

In order to analyze the possible confusion caused by the existence of secondary stress in P3 and P1 words, we looked into the reaction times of the participants (Table 19). The reaction time is the time counted from the moment every stimulus presentation ended to the moment the participant clicked on the desired button.

Table 19: reaction time (in seconds) average of the favorite positions per proficiency level per stress position in the perception test.

	P3	P2	P1
N1	3.4	3.6	3.8
N2	3.6	3.6	3.9
N3	3	3	4.1
N4	2.2	2.2	2.6

We can see from Table 19 that the reaction time tends to decrease with proficiency level increase, and also that P3 was always clicked first than P1 words. A hypothesis for such outcome regards order of stimulus presentation. That is, P3 words are by definition words where the first syllable is stressed. The immediate perception of prominence may have been enough for the participants to click on a certain syllable, regardless of its nature (primary or secondary). So, as soon as the participants perceived a stressed syllable, they would completely disregard other possibilities, and would get P3 words correct, which they did indeed (Figure 30).

If secondary stress posed a difficulty for the participants as we assume, it is intuitive that they would have chosen P1 for P3 words at times, and vice versa, since these two types of words alternate primary and secondary stress. We actually looked at that data, and a distribution of the chosen (favorite-FP) stress positions according to the position that was supposed to be clicked (correct-CP), per level, is presented in Figure 31 below.

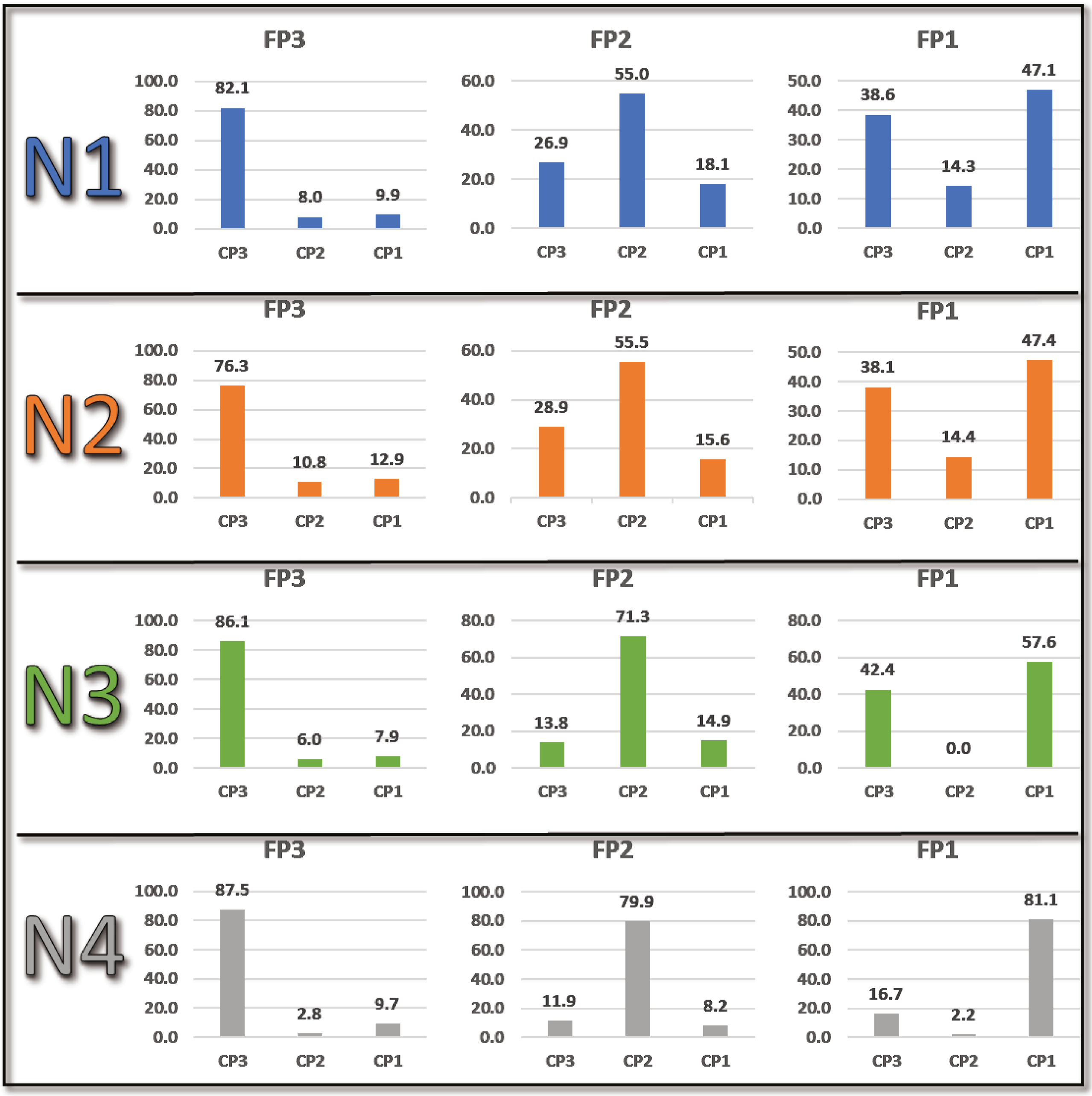


Figure 31: frequency (in %) of favorite stress positions (FP) per correct position (CP) per level (perception test).

For P3 words, participants of all levels almost always (~80% of the times) clicked on the right position. However, when observing their chosen positions for P1 (right-most part of the figure), there is almost an even distribution with P3! That is even more evident for N3 subjects, where their FP1s were exclusively either for P1 (58%) or P3 (42%). The scenario is about the same for N1 and N2, but for N4 speakers such pattern seemed to disappear.

The same does not happen for the P3 words. Most of the FP3 were actually a match with CP3, indicating no interference of secondary stress when clicking on the assumed correct stress position. Analyzing the acoustic data of the AE speaker used to create the experiment (Figure 32 below), we found no significant jumps from primarily to secondarily stressed syllables in any of the acoustic parameters that would justify the BP speakers to have chosen P3 correctly in their majority. We assume then, that the position (P3) had such high correct ratings because it is expected from words in English to start with a stressed syllable, as previously discussed in Chapter 5.

Figure 32: differences in 1ry-2ry stressed syllables of P1 and P3 words produced by the native AE speaker used in the perception test.

The P2 words were the easiest for all levels of proficiency. These words followed the pattern unstressed-stressed-unstressed syllables, without any competitive stimuli. The only real task was identifying stress instead of also having to differentiate its levels, making the task easier for BP speakers.

Overall, stress perception was an easy task for all participants in this study. However, differentiating words with initial and final stress based on the acoustic cues seemed tricky to all participants as well. What that tells us is that even though Brazilians are able to identify stress within words better than producing it (expected since both BP and AE are stress languages, therefore sensitive to the parameter), the acoustic features of primary and secondary stress are such that they cannot be distinguished by foreigners at a perception level.

8 Production and perception of lexical stress: relationship

We could observe that the BP speakers of all four levels of proficiency in English, as well as the AE speakers, employed mostly duration and relative intensity (correlate of vocal effort) to stress the syllable of the words from this study. The results are close to the ones presented by Barbosa et al. (2013), who also investigated the relevance of different acoustic parameters in BP to signal lexical stress in three different conditions: word reading, phrase reading, and informal interview. Duration was very distinctive for stress in all conditions, and stress reached an effect size up to 50% in the variance of the parameter, greater than F0 and intensity. The cognate status of the words effected changes in total intensity and peak of F0 only. We expected an influence of the cognate status in the acoustic parameters, but its major impact was in the correctness of stress placement and not on the acoustic realization of linguistic stress.

Cognate words that coincide in stress with BP (P1) had the most percentage of hits in the production test (average 85%, lowest 80%), whereas the ones with opposite stress (P3) had the poorest (average 43%, lowest 27%). This shows that when a word in a foreign language resembles a correspondent in L1, the prosodic aspects of L1 are triggered by them, leading to stress misplacement.

The perception experiments showed an improvement for the overall hits of all four levels, as well as the decrease in the influence of the cognate status over stress (differences C-NC diminished for all levels).

9 Accent Ratings

Audio samples of BP speakers were used to create two sets of accent rating tests. The presented stimuli were the same in each test. For this task, 10 native speakers of AE – some of whom were also part of the production test – listened to whole sentences of the BP speakers, N1-4, plus sentences from a native AE speaker used as a control, and had to judge each production using two different scales: Intelligibility and Nativeness.

9.1 Intelligibility

The AE raters were instructed to listen to the audio samples and give a score from 1-5 judging how easy to understand that particular audio was, 1 being poorly understandable and 5 completely understandable. Figure 33 below presents the percentage of scores per level in the category.

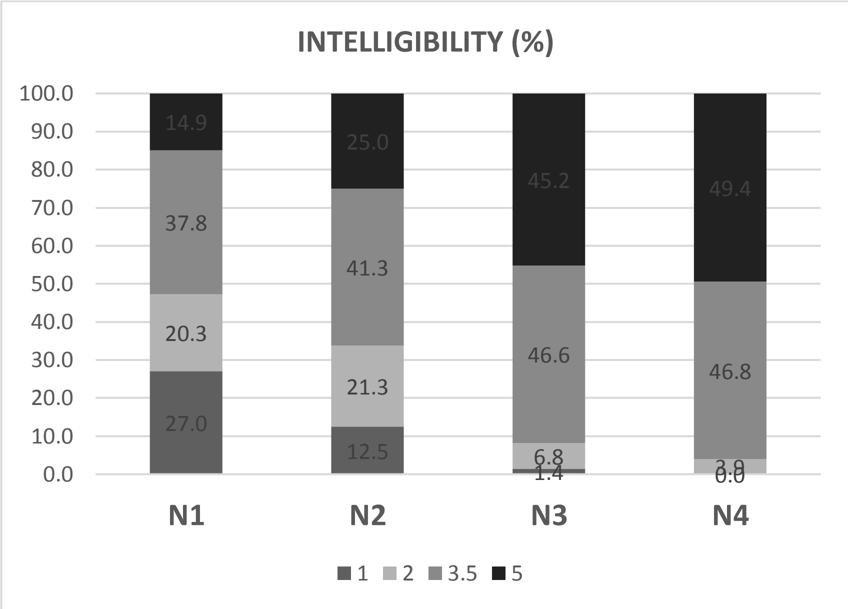


Figure 33: results for the accent rating – intelligibility.

The levels of proficiency seemed logically distributed along the ratings. The least proficient speakers (N1) had above half of their ratings within 1 and 2 scores, which means their production was poorly comprehended by native speakers of English, while N3-4 speakers had above half of their productions rated 5, extremely easy to understand.

The level of proficiency from the grammar/vocabulary test from the screening process seemed to go hand in hand with the pronunciation rated by native AE speakers. In this test, samples of one of the male AE speakers were mixed with the BP productions in order to see if the raters would give him a score different than 5, to control the rigorousness of the raters, but the stimulus received the highest score in 100% of the ratings.

9.2 Nativeness

For this scale, the same test was run, but the raters were asked to judge how close the productions were to a native production (Figure 34). It is foreseeable that the nativeness of the productions would be much more prone to receiving lower scores than intelligibility, since it is a more accent-related measure. In many productions – especially the ones from the N1-2 speakers – the message may have been understood but with difficulty, which may have caused a stimulus to be judged as mildly understandable but far from being pronounced by a native speaker of English.

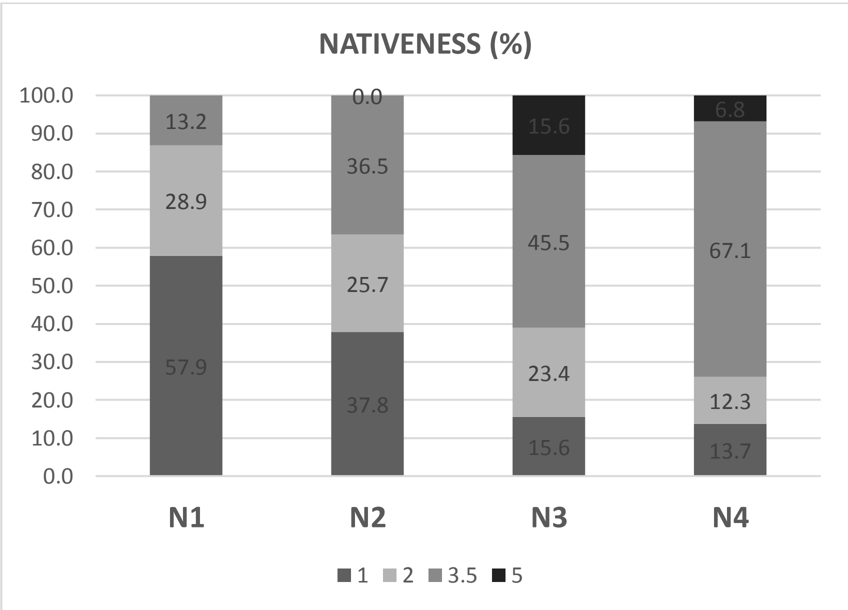


Figure 34: results for the accent rating – nativeness.

The picture shows that raters were much stricter rating the BP productions when judging nativeness. Only N3-4 speakers received the 5 score (a.k.a. were considered native like by AE speakers), and even then, it never was more often than 16% of the ratings. The N1 speakers had almost 60% of their ratings exclusively rated 1, being considered very nonnative by the AE raters, despite being considered well understood, according to the intelligibility scale.

Having a production judged as well understood, therefore containing the basic segmental elements of a sentence does not necessarily mean it will be considered native-like. This gap that separates the rigorousness between the Nativeness and Intelligibility scales relies on the specificities of the prosodic acoustic parameters before mentioned in this work, once again stepping on the important role prosody plays in the perception of a foreign accent, not ignoring that such accent may be perceived by mispronunciation of segments as well.

10 Final Remarks

For the sake of summarizing the main findings of this study, we hereby describe the most important outcomes and inferences that could be drawn from this study:

- Duration and relative intensity (vocal effort) showed to be important in the acoustic realization of lexical stress for both AE and BP speakers;

- It has been shown an inverse relationship for F0 parameters between the AE speakers who use the crest of the pitch contour to stress syllables, while the BP speakers use the trough;
- The cognate status of words influenced the production of TOTINT and F0PEAK only. Cognates influence the correctness of stress position rather than how the prominence is acoustic realized;
- Having BP as L1 showed to strongly influence stress placement in AE, since speakers scored an average of 85% of correct P1 words (where there is a stress match between both languages), but only 43% correct for P3 words (where stress completely differs between both languages).
- Speakers of all proficiency levels scored better in the LS perception test than in the production test, demonstrating that LS is easier to be perceived than to be produced (expected, since both languages in this study have LS, making them therefore sensitive to this feature). However, differentiating across stress levels is not such as an easy task. Words with secondary stress caused confusion amidst the BP speakers through levels N1-3.
- As for the accent ratings, Nativeness is much more resistant to be rated by native AE speakers than Intelligibility, even though both scales are coherent with the respective proficiency levels.

11 REFERENCES

ADANK P, SMITS R, van HOUT R, A comparison of vowel normalization procedures for language variation research. *J Acoust Soc Am*. Vol 116 (5). 2004.

ANDERSON-HSIEH J, Johnson R, Koehler, K. The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentals, Prosody, and Syllable Structure. *Language Learning*, 42. 1992.

ARAÚJO, G. A. et al. Algumas observações sobre as proparoxítonas e o sistema acentual do português. In: ARAÚJO, G. (Org.). *O acento em português: abordagens fonológicas*. São Paulo: Parábola, 2007

ASHER, JJ & García, R. 1969. The optimal age to learn a second language. *The Modern Language Journal*, 53.

BARBOSA, PA, Eriksson A, Åkesson J. On the robustness of some acoustic parameters for signalling word stress across styles in Brazilian Portuguese. *INTERSPEECH*. 2013.

BOERSMA, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.22, retrieved 07 December 2016 from http://www.praat.org/

CONSONI, F.; Barbosa, P. A.; Netto, W. F. 2006. Hipóteses acerca do reconhecimento de acento lexical em palavras isoladas. *Estudos Linguísticos* XXXV, 792-799.

COUNCIL OF EUROPE. Common European framework of reference for languages: Learning, teaching, assessment. Cambridge, U.K: Press Syndicate of the University of Cambridge. 2001.

CUTLER, A; Mehler, J. 1993. The periodicity bias. Journal of Phonetics, 21,103-108. Dupoux, E. et al. 1997. A distressing 'deafness'in French? *Journal of Memory and Language*, 36, 406-442.

DUNCAN, DB. Multiple range and multiple F tests. *Biometrics*. 11. 1955.

EXAM ENGLISH ltd. https://www.examenglish.com/leveltest/index.php . 2014.

FLEGE, J. 2018. It's input that matters most, not age. *Bilingualism: Language and Cognition*, 21(5), 919-920.

FLEGE, J., & MacKay, I. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition,* 26, 1–34.

FLEGE, JE. Second language speech learning: Theory, findings, and problems. In: W. Strange, (Ed.) *Speech perception and linguistic experience: issues in cross-language research*, Baltimore: York Press. 1995.

FLEGE, J. E., FRIEDA, E. M., NOZAWA, T. 1997. Amount of native-language (L1) use affects the pronunciation of an L2, *Journal of Phonetics*, 25, 169-186.

FLEGE, JE, Eefting, W. Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica* 43:155-71. 1986.

FLEGE, JE, and Hillenbrand, J. 1986. Differential use of temporal cues to the /s/-/z/ contrast by native and non-native speakers of English. Journal of the Acoustical Society of America 79:508-17.

FLEGE, JE., and Hillenbrand, J. 1987. Differential use of closure voicing and release burst as cue to stop voicing by native speakers of French and English. Journal of Phonetics 15:203-8.

JILKA, M. The contribution of intonation to the perception of foreign accent: Identifying intonational deviations by means of F0 generation and resynthesis. 2000, 206p. Thesis (Doctorate in Phylosophy) – Faculdade de Filosofia, Universidade de Stuttgart, Stuttgart, 2000.

KAGER, R. The Metrical Theory of Word Stress. In: Goldsmith, J.A. *The Handbook of Phonological Theory*. 1995.

KOCHANKSI, G. Prosody Beyond Fundamental Frequency. In: SUDHOFF, DL et al. *Methods in Empirical Prosody Research*. New York. 2006.

LEE, SH. Morfologia e fonologia lexical do português do Brasil. [Tese de Doutorado]. Campinas: Departamento de Linguística do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas. 1995.

LENNEBERG, E. 1967. Biological Foundations of Language. New York: Wiley.

LOBANOV, BM. Classification of Russian vowels spoken by different speakers. *J. Acoust. Soc.* Am. 49. 1971.

MAREÜIL, P. B.; Vieru-Dimulescu, B. The contribution of prosody to the perception of foreign accent. Phonetica, v. 63(4), p. 247-267, 2006

MASSINI-CAGLIARI, G. 1992. Acento e ritmo. São Paulo, Contexto. p.95.

MEHLER, J. et al. 1988. A precursor of language acquisition in young infants. Cognition, 29, 143-178.

MIYAWAKI, K., Liberman, A., Jenkins, J., and Fujimura, O. 1975. An effect of linguistic experience: The discrin'tination of Ir] and [l) by native speakers of Japanese and English. Perception & Psychophysics 18:331-40.

MOYER, A. 1999. Ultimate attainment in L2 phonology, Studies in Second Language Acquisition, 21. 81-108.

NAVARRO, F. R. Factores que influyen en el acento extranjero: estudio aplicado a aprendices estadounidenses de español. 2008, 356p. Tese (Doctorate in Spanish) – Faculdade de Filologia, Departamento de Língua Espanhola, Universidade de Salamanca, Salamanca, 2008.

PENFIELD, W. 1965. Conditioning the uncommitted cortex for language learning. Brain 88:787-98.

PISKE, T, MacKay, IRA, Flege, J. 2001. Factors affecting degree of foreign accent in an L2: a review. Journal of Phonetics, 29

PURCELL, ET., Suter, RW. 1980. Predictors of pronunciation accuracy: a reexamination. Language Learning, 30.

SUTER, RW. 1976. Predictors of pronunciation accuracy in second language learning. Language Learning. 26

SLUIJTER, A. M. C., Heuven, V. J., and Pacilly, J. J. A. 1997. "Spectral balance as a cue in the perception of linguistic stress," J. Acoust. Soc. Am. 1, 503–513.

TAHTA, S., Wood, M., Loewenthal, K. 1981. Foreign accents: factors relating to transfer of accent from the first language to a second language. Language and Speech, 13.

THOMPSON, I. 1991. Foreign accents revisited: the English pronunciation of Russian immigrants. Language Learning, 41.

TRAUNMÜLLER, H. and Eriksson, A.. 1995. The frequency range of the voice fundamental in the speech of male and female adults. Manuscrito inédito.

VOGEL, I. 2000. The acquisition of prosodic phonology: challenges for the L2 learner. Paper apresentado em Structure, Acquisition, and Change of Grammars: Phonological and Syntactic Aspects. Hamburg, Germany.

ZHANG, Y; Nissen, S. L.; Francis, A. L. 2008. Acoustic characteristics of English lexical stress produced by native Mandarin speakers J. Acoust. Soc. Am. 123(6).

ZONG, J. 2016. Frequently Requested Statistics on Immigrants and Immigration in the United States. Migration Policy Institute. Web. Nov 26. 2016

6 Appendixes

Appendix I- Sociolinguistic Questionnaire

Name						
Age						
Gender	M	F				
Where have you studied English? (High School, private sessions...)						
For how long you have studied English						
Have you ever lived in an English speaking country?				Y	N	
Self-evaluation of English skill level			1	2	3	4

Appendix II- Consent Forms

Informed Consent Form
PRODUCTION AND PERCEPTION OF WORD STRESS IN ENGLISH BY BRAZILIAN PORTUGUESE SPEAKERS
Filipe Modesto (advisee) and Plinio Almeida Barbos (advisor)
Master's Dissertation Project
IRB approval number: 66786617.2.0000.5404

You have been invited to participate voluntarily in this study. This is a document called Consent Form, that assures your rights and obligations as a participant, and there are two copies of it, one for you and one with the researcher.

Please, read it carefully, and clear any information that might have passed misunderstood. If there are any questions before or after you sign it, you can clear them with the researcher. You can take it to your family or significant others before you decide participating. If you decide not to or wish to withdraw from the study at any moment, there will be no punishment.

Rationale and goals: this study aims to understand how native speakers of Brazilian Portuguese process word stress in English, regarding its production and perception.

Methods: when agreeing to participate, you're invited to:

- Fill out a form with informations such as (name, ae…) and questions regarding your contact with foreign languages;
- Read a list of 73 words in English. This will be recorded for further acoustic analysis.
- Undergo an accent rating test for intelligibility and nativeness of some Brazilian speakers speaking English.

Discomfort and risks: as much as every research has its risks or discomforts, we consider the time required to set equipment ready as being a possible discomfort (approximately 25 minutes). There are no foreseeable risks.

Benefits: the study will not directly benefit the participant individually, but collectively, contributing to the experimentation in acoustic phonetics field.

Support: the researcher will be present during the entire time of the described procedures, for any required support, during and after the research is concluded. Your doubts will be cleared at any moment, if you have them.

Secrecy and privacy: you are assured secrecy about your information and data. No information will be released to others outside the scientific community, as well as the voice recordings. When publishing the outcomes of this research, your name will not be disclosed.

Refunding: there is no need for refunding, since the activities here mentioned will be previously scheduled.

Data Storing: the collected data of this research will be stored in a recording database at the Group of Speech Prosody Studies (GSPS), at the Institute of Language Studies of UNICAMP. It is a right of the participant to allow or not the use of the material in future studies. In this case, the material will only be used again after the IRB approval. Regarding your consent, check one of the following options:

[ ] I authorize the storing of the collected material in the database at the Group of Speech Prosody Studies (GSPS), at the Institute of Language Studies of UNICAMP

[ ] I authorize the use of the collected material in this research only after new approval by the IRB.

Contact: In case you have any questions about the study, please feel free to contact Filipe Modesto at: 571, Sergio Buarque de Holanda, Postal Code 13083-859, Campinas-SP, Brazil. Phone: +551935218936. Email: lipesssmodesto@gmail.com. In case of complaints or concerns about your participation in the study, you may also contact the Institutional Review Board (CEP, in Portuguese) at: 126, Tessalia de Vieira Camargo, Postal Code 13083-887, Campinas-SP, Brazil. Phone: +551935218936. Fax: +551935217187. Email: cep@fcm.unicamp.br

Informed Consent: after the presentation of this study, its objectives, methods, benefits, potential risks and discomforts, I hereby accept to participate:

(Name)____

(Signature)______________________________________(Date)______/______/_______.

Researcher responsibility: I hereby assure that I have fulfilled the requirements of the resolution 466/2012 CNS/MS when elaborating the protocol and of this Informed Consent Form. I also confirm that I explained every topic of this document and provided a copy of it to the participant. I inform that this study was approved by CEP and I will use this material and the collected data exclusively as presented or according to the participant consent.

(Researcher signature)______________________________________(Date)______/______/_______.

Appendix III- Approval Letter of the Ithaca College Institutional Review Board

ITHACA COLLEGE
Sponsored Research

January 17, 2018

Filipe Modesto
School of Health Sciences and Human Performance

Re: IRB 0118-09 - Lexical Stress in English by Brazilian Portuguese Speakers

The Institutional Review Board for Human Subjects Research (IRB) has reviewed your proposal and has determined that by the IRB Guidelines, this project can be approved for an exemption from ongoing oversight.

In certifying that your research is exempt, the IRB indicates that there will be no continued oversight. Should the project be continued beyond the semester, you may do so without additional involvement with IRB provided that the same procedures as described in the application are followed. If there are changes in design, the application would have to be resubmitted to the IRB.

College policy requires you to maintain, as part of your records, all correspondence with the IRB (including the complete, approved Request for Review or Application for Exemption), all documentation of informed consent, and any lists used in assigning codes or other identifiers to participants for a period of three years. Researchers may keep, at their discretion, completed data collection instruments provided they continue to be protected in the manner specified in the Request for Review or Application for Exemption and as described to the subjects in the process of obtaining informed consent.

This approval is issued under the Ithaca College's OHRP Federal-wide Assurance #00004870. Please feel free to contact the IRB at irb@ithaca.edu with any questions or concerns.

Best wishes for a successful study,

Sincerely,

Warren J. Calderone

Warren Calderone
Director of Corporate, Foundation Relations, and Sponsored Research
Institutional Review Board for Human Subjects Research

Appendix IV- Corpus: BRENGLISH-STRESS

WORD	COGNATE STATUS	COG. TYPE	NATIVE STRESS
abstinence	C	C-OTHER	P3
adventure	C	C-OTHER	P2
amplitude	C	C-OTHER	P3
authorize	C	C-OTHER	P3
basketball	C	C-OTHER	P3
critical	C	C-OTHER	P3
demonstrate	C	C-OTHER	P3
dignify	C	C-OTHER	P3
disappear	C	C-OTHER	P1
initial	C	C-OTHER	P2
interact	C	C-OTHER	P1
memorize	C	C-OTHER	P3
metaphor	C	C-OTHER	P3
musical	C	C-OTHER	P3
persistent	C	C-OTHER	P2
photograph	C	C-OTHER	P3
portuguese	C	C-OTHER	P1
positive	C	C-OTHER	P3
protestant	C	C-OTHER	P3
resistance	C	C-OTHER	P2
specify	C	C-OTHER	P3
summarize	C	C-OTHER	P3
sympathy	C	C-OTHER	P3
celebrate	C	Cate	P3
consummate	C	Cate	P3
decorate	C	Cate	P3
desperate	C	Cate	P3
suffocate	C	Cate	P3
disinfect	C	Cct	P1
disrespect	C	Cct	P1
introspect	C	Cct	P1
resurrect	C	Cct	P1
ambition	C	Ction	P2
audition	C	Ction	P2
confection	C	Ction	P2
position	C	Ction	P2
proportion	C	Ction	P2
acknowledge	NC	NC-OTHER	P2
allowance	NC	NC-OTHER	P2
appraisal	NC	NC-OTHER	P2
betrayal	NC	NC-OTHER	P2
challenger	NC	NC-OTHER	P3
dangerous	NC	NC-OTHER	P3
demeanor	NC	NC-OTHER	P2
edible	NC	NC-OTHER	P3
employer	NC	NC-OTHER	P2
gullible	NC	NC-OTHER	P3
mastery	NC	NC-OTHER	P3
meaningless	NC	NC-OTHER	P3
mispronounce	NC	NC-OTHER	P1
outbalance	NC	NC-OTHER	P2
overact	NC	NC-OTHER	P1
overprice	NC	NC-OTHER	P1
reachable	NC	NC-OTHER	P3
saturday	NC	NC-OTHER	P3
seasonal	NC	NC-OTHER	P3
spendable	NC	NC-OTHER	P3
standardize	NC	NC-OTHER	P3
underact	NC	NC-OTHER	P1
understand	NC	NC-OTHER	P1
abrogate	NC	NCate	P3
adumbrate	NC	NCate	P3
pullulate	NC	NCate	P3
stridulate	NC	NCate	P3
tessellate	NC	NCate	P3
benefact	NC	NCct	P1
introject	NC	NCct	P1
reconvict	NC	NCct	P1
apportion	NC	NCtion	P2
expunction	NC	NCtion	P2
misfunction	NC	NCtion	P2
sortition	NC	NCtion	P2
vacation	NC	NCtion	P2

*underlined syllables bear secondary stress