UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE LETRAS

BRUNO MORAES SCHWARTZHAUPT

# FACTORS INFLUENCING VOICE ONSET TIME: ANALYZING BRAZILIAN PORTUGUESE, ENGLISH AND INTERLANGUAGE DATA

Porto Alegre

2012

BRUNO MORAES SCHWARTZHAUPT

# FACTORS INFLUENCING VOICE ONSET TIME: ANALYZING BRAZILIAN PORTUGUESE, ENGLISH AND INTERLANGUAGE DATA

Monografia apresentada ao Curso de Letras da Universidade Federal do Rio Grande do Sul, como requisito parcial para obtenção do grau de Licenciado em Letras.

Orientador: Prof. Dr. Ubiratã Kickhöfel Alves

Porto Alegre
2012

If it’s worth doing,
it’s worth doing right.
(George Carlin)

## AGRADECIMENTOS

***(acknowledgements)***

Em primeiro lugar, é obrigação minha agradecer muito aos meus pais, Ercilio Schwartzhaupt e Rosita Moraes. Eu realmente acredito que, se cheguei a algum lugar, isso se deve a todo o apoio incondicional e os sacrifícios feitos por eles. Tenho neles exemplos de caráter a seguir, e muito me orgulharei se conseguir fazê-lo.

Devo agradecer, também, a todos os outros que fazem parte da minha família, tanto materna quanto paterna. Crescer num ambiente de tanta união, respeito e bom-humor foi um privilégio.

Não posso deixar de mencionar aqui os valiosos amigos que tenho; amigos que me proporcionam momentos sem os quais eu não teria paciência pra escrever tudo isto aqui: Rodrigo Ribeiro, Geisson Flesch, Davi Inácio e Alexandre 'Bento' Kuciak. Vocês sabem, gurizada: *whatever it takes,...* Vale um agradecimento especial pro Rodrigo, pela ajuda com as coletas; e outro pro Davi, que anteriormente trabalhou comigo em projeto similar a este.

Além disso, agradeço a Ubiratã Alves, meu amigo e excepcional orientador, com quem muito aprendi nos últimos anos, e ainda aprenderei muito mais, eu sei.

Também devo mencionar as duas professoras que integrarão a banca que avaliará este trabalho: Ingrid Finger e June Campos – profissionais que muito admiro, e com quem muito aprendi.

Sou grato a todos os participantes desta pesquisa, por terem sido tão pacientes.

Enfim, agradeço, também, aos bons amigos que encontrei na UFRGS, e àquelas pessoas que me motivaram em meu percurso, antes e especialmente agora, durante a escrita deste trabalho – pessoas incríveis, que tive o prazer de conhecer.

*Let us not get carried away.*

## ABSTRACT

The present study investigates the effects of place of articulation, vowel quality and number of syllables of the target word on the Voice Onset Time (VOT) of the voiceless stop consonants /p/, /t/ and /k/ in word initial position. Ten Brazilian proficient learners of English and five native speakers of American English participated in this study, which verified production in Brazilian Portuguese (L1), American English (L1) and Brazilian English (L2). Both groups of subjects participated in a production test with monosyllabic and disyllabic English words; a production test with disyllabic Brazilian Portuguese (BP) words was conducted with the Brazilian informants only. Target words were inserted in carrier sentences and measured in software Praat (BOERSMA & WEENINK, 2012) in order to verify their absolute and relative VOT values. Results show clear distinctions on VOT in each of the three places of articulation in American (L1) and Brazilian (L2) English productions, whereas productions in BP show no difference between VOT of bilabial and alveolar stop consonants. In BP and Brazilian English, analyses regarding vowel quality effects show that differences on mean VOT values can be related to vowel frontness, not only to vowel height, as suggested by previous studies; furthermore, it could be observed that vowel quality effects vary according to the place of articulation of the stop segment. The same vowel quality effects could not be confirmed in productions in American English (L1). No effects concerning the number of syllables of the target word were verified in any of the three language systems. Overall, results show that place of articulation and vowel quality play a fundamental role in determining VOT values.



**Keywords:** Voice Onset Time, place of articulation, vowel quality.

## RESUMO

O presente estudo investiga os efeitos de ponto de articulação, qualidade da vogal e número de sílabas da palavra-alvo no intervalo de Voice Onset Time (VOT) das consoantes plosivas surdas /p/, /t/ e /k/ em posição inicial de palavra. Dez aprendizes Brasileiros de Inglês em nível proficiente e cinco falantes nativos de Inglês Americano participaram deste estudo, que verificou produções em Português Brasileiro (L1), Inglês Americano (L1) e Inglês produzido por Brasileiros (L2). Ambos os grupos de participantes foram submetidos a um teste de produção com palavras monossilábicas e dissilábicas do Inglês; um teste de produção com palavras dissilábicas do Português Brasileiro (PB) foi conduzido somente com os participantes Brasileiros. As palavras-alvo foram inseridas em frases-veículo e medidas, com o auxílio do *software* Praat (BOERSMA & WEENINK, 2012), quanto aos seus valores absolutos e relativos de VOT. Os resultados mostram distinções claras entre o VOT de cada um dos três pontos de articulação em produções em Inglês Americano (L1) e Inglês produzido por Brasileiros (L2), enquanto produções no PB não revelam diferenças entre o VOT de plosivas bilabiais e alveolares. Tanto no PB quanto no Inglês produzido por Brasileiros, análises quanto aos efeitos da qualidade da vogal mostram que diferenças nas médias de VOT podem estar relacionadas à anterioridade da vogal, e não apenas à altura da mesma, como sugerido por estudos anteriores; além disso, pode-se observar que os efeitos quanto à qualidade da vogal variam de acordo com o ponto de articulação da plosiva. Tais efeitos não são observados, no entanto, em produções em Inglês Americano (L1). Efeitos quanto ao número de sílabas da palavra-alvo não foram observados em nenhum dos três sistemas linguísticos investigados. Em geral, os resultados mostram que o ponto de articulação da plosiva e a qualidade da vogal seguinte desempenham papéis fundamentais na determinação de valores de VOT.



**Palavras-chave:** Voice Onset Time, ponto de articulação, qualidade da vogal.

## LIST OF BOXES

## LIST OF FIGURES

## LIST OF TABLES

**CONTENTS**

## 1 INTRODUCTION

In the last few years, a great amount of the studies directed to the phonetic-phonological acquisition of English as a Second Language[1] (hereafter L2) analyzed perception and production of *Voice Onset Time* (hereafter VOT) patterns. A large variety of works can be cited in the recent Brazilian scenario (COHEN, 2004; ZIMMER, 2004; ALVES, 2007; REIS & NOBRE-OLIVEIRA, 2008; GEWEHR-BORELLA, 2010; MAGRO, 2010; ALVES & MAGRO, 2011; ALVES *et al*, 2011; FRANÇA, 2011; M. ALVES, 2011; PRESTES, 2012). The investigation of this *phonetic-phonological*[2] aspect is of great importance due to its distinctive perceptual character between voiced and voiceless stops not only in English, but also in many other world languages.

The recently facilitated access to important investigation tools, such as software *Praat* (BOERSMA & WEENINK, 2012)[3], certainly favored academic production involving acoustic analysis. Another reason for researchers to often draw their attention to VOT lies in the fact that they find it a rich object of study, which allows them to characterize voicing distinctions in stop consonants in a gradient fashion, and goes beyond the binary 'voiceless' *versus* 'voiced' contrast. An acoustic VOT analysis transposes linguistic data into numbers, making it possible for linguists to verify the acquisition of this phonetic-phonological aspect more precisely. Through an acoustic VOT analysis, linguists can observe, for example, objective data on developmental stages of L2 acquisition, when the VOT values produced by a particular learner are intermediate to the patterns found in the learner's mother tongue and the ones found in the target language.

VOT values are not absolute; they are rather influenced by several different factors. Considering that VOT is such a valuable object of study, it is essential that researchers have a clear understanding of how it is characterized and what variations are expected when dealing with this phonetic-phonological aspect. Therefore, in order to have a more reliable method, and avoid misinterpretation of results, research studies should take factors influencing VOT into consideration. This is of great importance as, for instance, making a thorough selection of

[1] The present study does not distinguish between the terms *Second Language* and *Foreign Language*. The former shall be used here as a synonym of the latter.

[2] The term *phonetic-phonological* is used in the present study, rather than *phonetic* or *phonological*, as we conceive the existence of a gradient, a continuum between the physical *phone* and the *phoneme* (ALBANO, 2001, 2002; ZIMMER & ALVES, 2006).

[3] This free software allows its users, among other functions, to perform acoustic verification and to elaborate perception tests. Further information and download at [http://www.fon.hum.uva.nl/praat/].

target words based on the contexts in which the plosive consonants are presented is crucial for investigators to understand resultant VOT values in production tests or accuracy levels in perception tests.

Not much research has been conducted, however, aiming to investigate such factors influencing VOT. Cho & Ladefoged (1999) tested universals and variations on VOT across 18 endangered languages, which varied widely concerning voicing patterns. Although their primary objective was not to test place of articulation effects, the authors presented data supporting the idea that, as to bilabial, alveolar and velar stop consonants, the latter one was produced with longer VOT in the major part of the languages which contrasted stops in at least two of these three places of articulation[4].

Yavas & Wildermuth (2006) and Yavas (2008) investigated the effects of place of articulation of the stop consonant and the height of the vowel that follows this consonant on VOT in the L2 production of English by native speakers of Spanish. The authors found that these phonetic-phonological contexts are among the factors which have a significant influence in VOT values: the farther back the place of articulation of the plosive, the longer the VOT is going to be; also, C + V sequences with high vowels tend to show longer VOT than those with low ones. Taking these works in consideration, França (2011) also tested vowel-height and place of articulation effects on VOT produced by Brazilian learners of English in their L1 and L2, having similar findings. Effects regarding the number of syllables of the target words were also found by França (2011), who verified a tendency of longer VOT values in monosyllabic target words.

Nonetheless, as it shall be discussed further, there are still some considerations to be made concerning such effects, despite the results obtained in these studies. Firstly, it is important to pursue this verification, by testing not only native production of different L1 systems but how these systems interact in *interlanguage*[5] production. This can provide a more comprehensive view of this phenomenon as, by conducting analyses in these circumstances, a researcher can establish comparisons between the effects observed in each of the linguistic systems involved, as well as the possible influence one can have over the other in L2 acquisition.

Furthermore, it is still necessary to test the role played by vowels as to their frontness. As mentioned above, studies were conducted to verify vowel height effects on VOT; no

[4] It is important to realize that many of these languages did not have voiceless stops in the three places of articulation here in question (bilabial, alveolar and velar) in their phonological inventory.

[5] The definition for this concept is presented in subsection 2.5.

comparisons have been established, however, between front and back vowels of the same height. Since the effects are explained based on differences in the articulation of these vowels, it should be important to test differences between front and back vowels, which differ in terms of articulation.

The present study, therefore, aims to complement the abovementioned findings: effects of place of articulation (bilabial, alveolar, and velar), number of syllables (monosyllabic and disyllabic), and vowel quality[6] (high-front, low-front, high-back, and low-back) on the VOT of voiceless stop consonants were tested. In order for us to have a more comprehensive analysis, Brazilian Portuguese (hereafter BP), English and *interlanguage* data were verified. Moreover, as further explained in subsection 3.7, two measurements were used in this analysis: absolute and relative VOT values – the verification of divergences in the statistical significance of contrasts with both measurements, which has not been conducted so far, is another contribution brought by this study.

This study is composed of the following sections: (1) *Introduction*; (2) *Theoretical Background*, in which a review of the literature underlying the present study is made; (3) *Method*, in which the goal of this study and the methodological procedures used for conducting the data collection and the analysis are described; (4) *Results and Discussion*, which presents the results for the experiment conducted in this study and their implications; and (5) *Conclusion*, in which considerations are made regarding the findings and limitations of the present study.

The findings of the present study are expected to deepen the understanding of the factors influencing VOT variation in different linguistic systems.

[6] As explained, the literature has traditionally analyzed this distinction only by taking vowel-height effects into consideration. Since the present study also analyzes vowel-frontness effects, the term 'vowel quality' is going to be used in this text.

# 2 THEORETICAL BACKGROUND

In this section, a review of the literature underlying the present study is made. Firstly (2.1), the conception of phonetic-phonological acquisition which bases this investigation is presented. The next subsection (2.2) discusses the concepts of *phonetic-phonological transfer* and *grapho-phonic-phonological transfer*, which are fundamental for understanding how the L2 phonetic-phonological aspect in question is perceived and produced by Brazilian learners. The following subsection (2.3) explains the concept of *Voice Onset Time*. Subsection 2.4 discusses the documented effects of place of articulation, vowel height and the number of syllables of the target word on VOT. Finally, in subsection 2.5, the implications of the findings from previous studies which investigated the factors influencing VOT are discussed are discussed.

## *2.1 A dynamic phonetic-phonological acquisition perspective*

This section briefly presents the main tenets of the emergentist view of phonetic-phonological acquisition that guide the present study. Although what mainly motivates the development of this study is not primarily an investigation on the acquisition of *positive VOT* patterns by Brazilian learners, L2 production and the interaction between learners' L1 and L2 systems are analyzed here. Therefore, it is important to state what conception of phonetic-phonological acquisition underlies the ideas presented in this research study.

According to the emergentist view, once engaged in the acquisition of an L2, learners undergo a process which, as well as any other cognitive process, is *dynamic*. This means that several different factors, linguistic and non-linguistic, have influence over their learning process, and these factors cannot be considered in an isolated manner (DE BOT *et al.*, 2007).

In such a perspective, the linguistic input is considered rich, and it is responsible for guiding L2 acquisition: the input to which learners are exposed drives their performances and, based on the constraints and regularities presented by this input, the learners' awareness about the patterns of the target language system is raised (ZIMMER, SILVEIRA & ALVES, 2009).

L2 speech perception, a part of this dynamic process, occurs in a *multimodal* manner: multiple acoustic cues determine perception of segments, and these cues are not perceived by the learner in an isolated way (ZIMMER, SILVEIRA & ALVES, 2009; ZIMMER & ALVES, 2012). Moreover, certain acoustic clues may not play the same relevant role in different L1 systems – in some cases, in order to acquire an L2 phonetic-phonological aspect, learners

must perceive an acoustic cue which is not relevant in their L1 system, which makes this process even more difficult.

Furthermore, as explained by Zimmer & Alves (2008, 2010), oral L2 production also deals with the orchestration of multiple acoustic cues, which act together as a whole. This process comprehends the physical and abstract levels, which also go beyond binary perspectives. In other words, in the acquisition of an L2, learners have to perceive, and then produce, acoustic cues which are relevant for the target system.

### *2.2 L2 production: phonetic-phonological and grapho-phonic-phonological L1-L2 transfer*

The *Speech Learning Model* (FLEGE, 1995) and the *Perceptual Assimilation Model – L2* (BEST & TYLER, 2007) attempt to explain the segmental phonetic-phonological acquisition phenomenon of transfer between L1 and L2 knowledge. This study is fundamentally based on the model proposed by Best & Tyler (*op cit*), for this is more compatible with the multimodal phonetic-phonological acquisition conception underlying the present study (discussed in the previous section).

According to Best & Tyler (2007), since the phonic elements of the learner's L1 and L2 systems interact in a common phonological space, the L2 learner tends not to perceive which acoustic categories belong to their L1 and which belong to the L2 in question. This is to say that, once learners are faced with a "new" L2 sound, they might not create a new category for that sound. The assumption is that, instead, learners assimilate the new sound to the L1 pattern, by following their L1 acoustic cues, thus considering it as an already existing sound from their L1 phonological space.

This premise allows one to explain the difficulties found in the acquisition of *positive VOT* (aspiration) by Brazilian learners in the following manner: without formal instruction, these L2 learners tend not to perceive the divergences existent between the BP and English voicing patterns in stop consonants production. Consequently, as *positive VOT* (aspiration)[7] is not a relevant acoustic clue in their L1 system, learners assimilate this pattern to the one from BP (with unaspirated plosive segments) and, therefore, do not produce the target aspiration.

Another problem faced by L2 learners in the acquisition of the phonetic-phonological aspect in question is pointed out by Zimmer, Silveira & Alves (2009). Such difficulty lies in the fact that BP and English, in spite of making use of the same alphabetical system, follow

[7] As it shall be explained in the next subsection, voiceless stops /p/ /t/ and /k/ are produced with *positive VOT* (aspiration) in English, and with *zero VOT* (no aspiration) in BP.

considerably different patterns concerning the relationship established between orthography and sound. More specifically, the *grapho-phonic-phonological*[8] relation in BP is transparent (orthography tends to represent pronunciation more clearly), whereas this relationship in English is much more opaque. As a consequence of the entrenched knowledge from their L1, learners tend to transfer the grapho-phonic-phonological patterns to their oral production in the L2 (ZIMMER & ALVES, 2006).

As regards the acquisition of *positive VOT* by Brazilians, grapho-phonic-phonological transfer is a factor which reinforces the lack of assimilation of acoustic cues from the target language: considering that the graphemes 'p', 't' and 'k' correspond to *zero VOT* stop consonants in the learner's L1 sound system, in their L2 oral production, this learner tends to associate the sounds represented by these graphemes in the target language (aspirated) to the ones they would represent in their mother tongue (unaspirated).

This is consonant with the multimodal conception of phonetic-phonological acquisition presented in the previous subsection: both the acoustic and the orthographic stimuli (different sources of L2 input) can either work to oppose or to reinforce one another. Once learners assimilate L2 voicing patterns in accordance with their L1 knowledge, the orthographic stimulus may then be considered a source of reinforcement of the L1 pattern. If no assimilation occurred, it could be stated, both sources of input could be in competition, as the former would instantiate the L2 target forms, whereas the latter could be reinforcing the L1 pattern.

Therefore, one must observe, when considering the acquistion of English *positive VOT* by Brazilian learners, that it might be impossible to consider the phonetic-phonological or the grapho-phonic-phonological transfer separately on theoretical grounds. Within a multimodal phonetic-phonological acquisition perspective, these factors may contribute in equal and simultaneous ways for Brazilian learners not to acquire the L2 voicing pattern.

Recent studies investigating the acquisition of English *positive VOT* by Brazilian learners have shown, however, that proficient learners' productions may actually be close to the target language pattern, especially in relation to the velar consonant /k/ (REIS E NOBRE-OLIVEIRA, 2008; ALVES *et al*, 2011; FRANÇA, 2011; M. ALVES, 2011). In this case, considering the velar stop consonant, a reasonable explanation lies in the fact that, as it will be

[8] Zimmer & Alves (2006) describe this relation as *grapho-phonic-phonological* as an indication of the existence of a relationship between the orthographic form and the phones of the linguistic system in question. In this perspective, the traditional concepts of *phone* and *phoneme* correspond to a single reality. The authors (*op cit*) believe that the use of this term is successful in expressing this relationship, for such a term, in this conception, does not refer to unities of a symbolic nature.

addressed in the section which follows, VOT values found for that place of articulation in Southern BP are already distant from what can be considered a *zero VOT* pattern. This would be to say that Brazilian learners, or at least those with a Southern variety of BP, might have their acquisition of English *positive VOT* facilitated as regards velar plosive segments.

### *2.3 Voice Onset Time: explaining the phonetic-phonological aspect in question*

Stop consonants (/p/, /b/, /t/, /d/, /k/, /g/) are produced in three subsequent stages: at first, the involved articulators produce a total obstruction of the air in the oral cavity (closure); following this closure, there is the consonant release, which corresponds to the articulatory phase in which the obstruction is undone; finally, the vocal cords of the segment which follows this consonant start vibrating (KENT & READ, 2002; ASHBY & MAIDMENT, 2005;). VOT refers to the period of time between the stop consonant release and the vibration of the vocal cords of the segment following this consonant. Three main VOT patterns can be found in the languages of the world (LISKER & ABRAMSON, 1964; COHEN, 2004; REIS & NOBRE-OLIVEIRA, 2008):

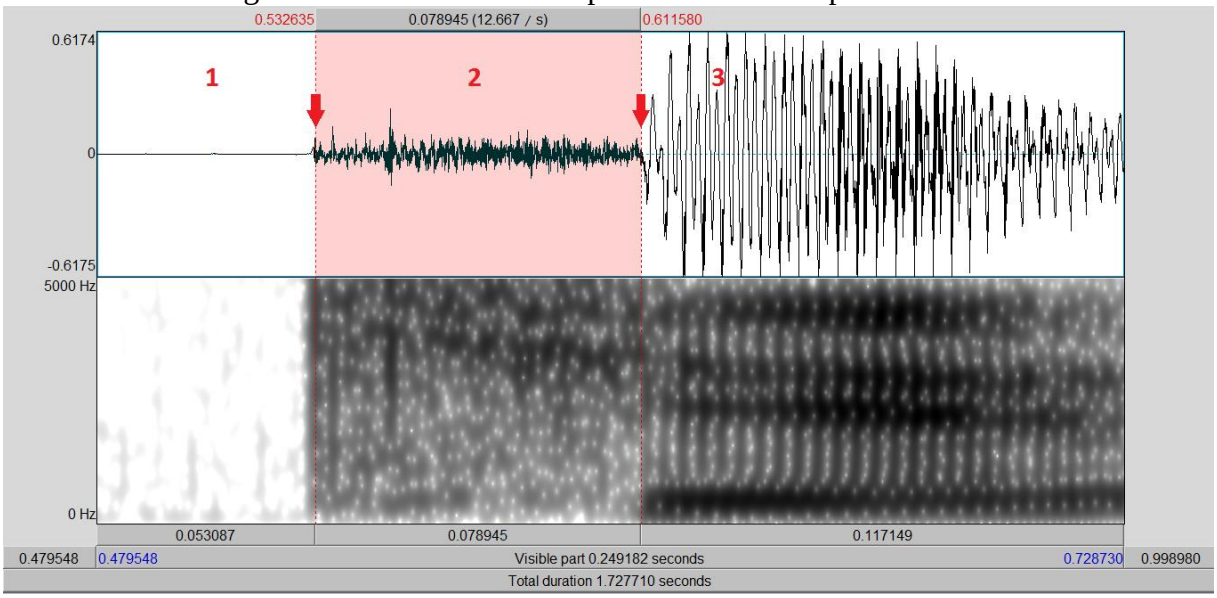
- ***Negative VOT*** (pre-voicing): the vocal cords start vibrating before the stop consonant release, in an interval ranging from -125 ms to -75ms;

- ***Zero VOT***: the vibration of the vocal cords starts almost simultaneously to the plosive release, in an interval ranging from 0 ms to +35 ms;

- ***Positive VOT*** (*aspiration)*: a delay follows the plosive release, and vocal cords start vibrating after a 35 ms to 100 ms interval.

*Figure 1*, which follows, illustrates the wave form and the spectrogram which characterize the production of the velar /k/ stop consonant in the word "*kit*" (preceded by a pause), produced by a native speaker of American English, as viewed in software *Praat* (BOERSMA & WEENINK, 2012). *Number 1* corresponds to the closure period (in which there is only silence), *number 2* stands for VOT (78,94 ms), and *number 3* corresponds to the [ɪ] vowel. The first arrow (on the left) signalizes the moment when the plosive is released; the second arrow marks the moment at which the vocal cords start vibrating.

Figure 1 – Illustration of the production of a stop consonant.




*Source: author's databank.*

In accordance with the literature cited above, BP voiced stop consonants /b/, /d/ and /g/ are produced with *negative VOT*, whereas voiceless plosives are produced with *zero VOT*, with mean values of approximately 12 ms for /p/, 18 ms for /t/ and 38 ms for /k/. Nevertheless, recent studies investigating the production of stop segments in the Southern region of Brazil have shown higher VOT values, especially for the velar stop /k/: Reis & Nobre-Oliveira (2008) found mean VOT values of 46.55 ms for this plosive; Gewehr-Borella (2010) measured mean VOT values for /k/ in two cities, verifying 52.43 ms and 63.90 ms[9]; the findings in França (2011) reveal mean VOT values of 47.20 ms for the same consonant. Such findings might indicate the existence of *partial aspiration* of voiceless /k/ in Southern Brazilian Portuguese.

As regards the production of stop consonants in English, voiced plosives tend to be produced with *zero VOT* (although productions with *negative VOT* may occur). Voiceless stops, on the other hand, are produced with positive VOT: [$p^h$] with mean 55 ms, [$t^h$] with mean 70ms, and [$k^h$] with average 80 ms VOT. Considering the existing divergences between BP and English voicing patterns in plosive segments, the two languages belong to distinct groups concerning VOT patterns.

It is essential to notice, however, as addressed by Cohen (2004), Reis & Nobre-Oliveira (2008), and Alves (2010), that VOT values are not absolute. VOT cannot be

[9] The former mean VOT value for /k/ was verified in *Rio Grande*, and the latter in *Picada Café*, both cities from the Brazilian Southernmost state, Rio Grande do Sul. The author analyzed production of monolingual and bilingual (BP + Hunsrückisch) participants, the values presented here were found in segments produced by monolingual participants.

considered as an isolated entity in any linguistic system. Several factors, which deserve consideration, influence this phonetic-phonological aspect, which deserve consideration. Among these factors, syllable stress, prosody, speech rate and the factors investigated by the present study (place of articulation, number of syllables and vowel quality) should be taken into account.

Since VOT values vary widely, it is reasonable for researchers to look for alternatives to have a more precise measurement of VOT, trying to eradicate possible misinterpretation of the acoustic data due to influence from the abovementioned factors. Traditionally, studies tend to measure absolute VOT values per se from isolated words. A different approach can be found in Prestes (2012)[10]: target words with initial stops were inserted in sentences, which would have their entire length measured. Both VOT length and sentence length were calculated with the objective to determine what percentage of the sentence was occupied by VOT (that is, the *relative value* of VOT in the target sentence).

This measurement may be a solution when researchers face inter-speaker variation concerning speech rate, for example: every participant shall have their own speech rate while reading the sentences aloud; clearly, higher speech rates tend to present lower VOT values. In this case, if absolute VOT values are measured and described, differences might not be related to the presence or absence of aspiration, but mostly to the speech rate with which the target sentence was read. Prestes (2012) could still verify place of articulation effects by comparing relative value means. Analyzing data from *relative values* (hereafter RV) seems, thus, to be a satisfactory alternative.

Considering these points, as explained further in section 3.7, a methodological choice was made in this study, which was to make use of target words in sentences and present both absolute and relative VOT values within the sentences. As explained in the introduction, although this is not initially a research question, verifying possible divergences originated from the difference in measurements shall thus be another contribution of the present study.

[10] The author analyzed production of English target words by Native Speakers of American English and Brazilian Learners.

### *2.4 The documented effects of place of articulation, vowel quality and number of syllables on VOT and their implications*

In this section we present what was found by previous studies which verified the effects of place of articulation, the quality of the subsequent vowel and the number of syllables of the target word on Voice Onset Time.

Cho & Ladefoged (1999) analyzed production in 18 endangered languages, and documented comparisons of mean VOT values in plosive segments with different places of articulation. The authors explain variation in VOT as to place of articulation effects based on *transglottal* pressure: both supraglottal and subglottal cavities are zones which affect the degree of abruptness of the pressure drop in the production of a stop consonant. Pressure in the supraglottal cavity has to be lower than pressure in the subglottal cavity for voicing to be produced. Two points of view are thus presented: firstly, since the supraglottal cavity formed behind the point of constriction in the production of velar stop consonants is smaller than the one formed in alveolar and bilabial stops, there is a higher pressure to be dropped before voicing can be produced – therefore, VOT (which corresponds to a period of voicelessness) is longer. Secondly, comparing the three places of articulation, the velar plosive has the largest body of air in front of the point of constriction. This body of air is considered as a mass which has to be moved before the air coming from the subglottal cavity can flow – therefore, once again, the constriction in velar stops slows the voicing lead.

Yavas & Wildermuth (2006) were the first to test the effects of place of articulation and vowel height on VOT in L2 production: the authors conducted a study investigating the production of the stop consonants /p/, /t/ and /k/ in English words as produced by 20 Spanish speakers of English as a Second Language. This study was conducted based on the premise found in Cho & Ladefoged (*op cit*) that a higher degree of abruptness of the pressure drop in the production of a stop consonant should make the voicing lead of the following segment start sooner.

In other words, the faster the articulators involved in the production of the plosive move farther away from each other, the shorter the VOT will be. That happens because the obstruction of the oral cavity is thus more rapidly undone, which causes the degree of the pressure drop in the oral cavity to be higher. That being the case, the following explanation can be offered: comparing bilabial, alveolar and velar plosive segments, the latter one is going to have the longest VOT, since the active articulator involved in the velar plosive /k/

production is the tongue dorsum, which moves slower than the tongue tip (used for articulating the alveolar plosive /t/) and the lips (used for articulating the bilabial plosive /b/).

As regards vowel height effects, Yavas & Wildermuth (2006) state that plosive segments should have longer VOT values when followed by high vowels. The explanation is given in relation to the opening of the oral cavity in the production of the subsequent vowels. As the authors explain, "since the high tongue position that is assumed during the stop closure in anticipation of a subsequent high vowel would result in a less abrupt pressure drop, a stop produced as such will have a longer lag than before a low vowel" (YAVAS & WILDERMUTH, 2006, p.253).

The findings in Yavas & Wildermuth (*op cit*) corroborated their hypotheses (discussed in the previous paragraphs); indicating, in /p/ (17 ms for low and 32 ms for high vowels) a shorter VOT than in /t/ (42 ms for low and 59 ms for high vowels); in turn, /t/ had shorter VOT than /k/ (51 ms for low and 89 ms for high vowels). Their results, therefore, suggest that, indeed, plosive segments followed by high vowels present longer VOT values than those followed by low vowels. In 2008, Yavas published another article with the same goals, and his previous findings were reassured.

The explanations provided by Yavas & Wildermuth (2006) refer to a matter of aerodynamics rather than a phenomenon of the linguistic system in question. In accordance with what the authors state, therefore, the place of articulation and vowel height effects on VOT explained above are *universal* rather than *language-specific*; they are characteristic of the human speech and are going to be verified in any language investigated.

As addressed earlier, França (2011) obtained production data from Brazilian speakers in BP and English, based on what was established by the investigation developed in the studies cited above, which were again corroborated. It is interesting to notice, however, that the results presented by França (*op cit*) for BP stops production show that, although all three contrasts were statistically significant, /p/ (19.56 ms) and /t/ (21.66 ms) presented similar mean VOT values, whereas /k/ (47.20 ms) is highly deviant.

As for the effects caused by the number of syllables of the target word on VOT, França (2011) verified, in 22 Brazilian learners' production of English monosyllabic and disyllabic words, that monosyllabic words tend to show slightly longer VOT values. This effect is more evident when comparing target words initiated by the alveolar stop consonant /t/: the mean VOT value in monosyllabic target words was 50.28 ms, whereas, in disyllabic words, it was 41.39 ms.

There are, nevertheless, some considerations to be made in relation to the methodology with which these findings were achieved and their real implications. Grounded on a multimodal phonetic-phonological acquisition perspective, the next subsection intends to discuss the findings presented above.

### *2.5 Considerations on previous studies on the factors influencing VOT and the gaps to be filled in*

Some considerations must be made in respect to the studies mentioned in the previous subsection. Yavas & Wildermuth (2006) and Yavas (2008) analyzed L2 English production of native speakers with very different language backgrounds. In relation to Yavas & Wildermuth (2006), although, officially, the L1 of all of these participants (20) was Spanish, they had been born in several different countries[11], and some of them had been living in the United States for approximately five years.

Once we observe that VOT variation within languages may be wide[12], it is reasonable to expect that these participants, who spoke different varieties of Spanish, produced stop consonants with considerably different VOT values. Such possibility was not, however, investigated by the authors: participants were tested in their L2 production only. It would be also important to understand whether the VOT values verified in the stop segments produced by these learners in L2 were based on the L2 input or their L1 variety of Spanish.

Even though the influence exerted by the factors here investigated on VOT is expected to be universal rather than language-specific, for they are related to aerodynamics, it is necessary to verify production not only in different L1 systems, but also among different varieties of the same language. The comparison of these effects in different L1 systems (as proposed by the present study) may provide strong evidence of their universality or, on the other hand, refute such a hypothesis. Moreover, to analyze *interlanguage*[13] data, in this sense, is to provide significant contribution to the understanding of similarities among language

[11] As expressed by the authors, participants had "*the following countries of origin: Cuba (5 subjects), Venezuela 5, Columbia 4, Peru 4, Costa Rica 1, and Mexico 1"* (YAVAS & WILDERMUTH, 2006).

[12] Cho & Ladefoged (1999), for instance, have categorized varieties of the same languages as 'aspirated' and 'unaspirated'. Also, as discussed in the previous sections, there is suggestion that Southern BP has longer VOT in comparison to other varieties of BP across the country.

[13] This work conceives *interlanguage* as a linguistic system shown by an L2 learner, which has its own structure, and is composed not only by aspects of the source and target language, but also by elements which may not be verified in their L1 or in the target language (GASS & SELINKER, 2008).

systems, as the interaction of two language systems in a new system with its own characteristics is tested – in which, following the hypothesis which has been assumed here[14], the same effects should be verified.

Another point to be observed is the fact that Yavas & Wildermuth (2006) and Yavas (2008) did not test production of stop segments followed by back vowels. Especially because the effects of the subsequent vowel on VOT are related to articulation, it is important to compare production of front and back vowels in the same height (as proposed by the present study). Many differences in the articulation of front and back vowels can be pointed out: front vowels are produced with the front part of the tongue, forming a large body of air in the back of the oral cavity and small one behind the front teeth. Back vowels, however, have very different tongue shapes, being produced with the back part of the tongue and originating two larger bodies of air, one in the throat and another one in the front part of the mouth (LADEFOGED, 2005)[15]. It is reasonable to assume, therefore, that these differences in the articulation of front and back vowels have an effect on the VOT of the preceding plosive.

Even though the investigation conducted by França (2011) tested its participant's L1 (BP) and L2 (English) production, no analysis comparing mean VOT values as to vowel frontness was carried out. This is, once again, an analysis which still has to be done in order to deepen the understanding of vowel quality effects on VOT.

Summarizing this subsection, it can be seen that some gaps were left by the previous investigations mentioned here, which the present study intends to fill in, such as the verification of vowel frontness effects on VOT, and the interaction between two language systems as regards these effects. Both analyses allow us to discuss similarities and differences across language systems, as well as the role played by aerodynamics in these effects.

[14] That is, the hypothesis established by Yavas & Wildermuth (2006) that the effects of the factors place of articulation and vowel height on VOT are aerodynamically universal rather than language-specific.

[15] These differences in size and position of the bodies of air formed in the oral cavity during the production of vowels, as explained by the author, are responsible for the rising or lowering of the first and second vowel formant frequencies (F1 and F2). For further information, see Ladefoged (2005).

# 3 METHOD

This section details the methodology based on which this research study was carried out. Firstly, the goals of this study are presented, and the hypotheses originated from them are explained. Information on the participants is presented in the following subsection. Next, the process of selection of target words is explained in two different subsections, one for English and another one for BP. The subsequent section contains information about the production test whose data correspond to the object of the analyses conducted in this study. Finally, the last section details how the acoustic verification and the statistical analysis were performed.

## *3.1 Goals of this study*

The present research study aims to verify, using two different measurements[16], some of the factors influencing VOT, namely place of articulation, number of syllables and vowel quality, in the production data of three different language systems: native Brazilian Portuguese, native American English and Brazilian proficient learners' English. In order to test these factors, the hypotheses expressed in the next subsection were established.

## *3.2 Hypotheses*

The hypotheses to be tested in this study were established concerning each of the three language systems investigated. No divergence in results is expected to be found in relation to the way in which the phonetic-phonological aspect in question was measured (absolute or relative values): hypotheses are the same for tests with both measurements. No hypothesis regarding the number of syllables was established for BP, since, as it shall be explained in subsection 3.5, the BP words investigated are all monosyllabic.

### 3.2.1 L1 Brazilian Portuguese Production

a) There are going to be significant differences in both absolute and relative VOT values concerning the place of articulation of the stop consonant. Bilabial /p/ and

[16] Absolute and relative values of VOT within a target sentence, as further explained in subsection 3.7.

alveolar /t/ are not going to be significantly different from one another; the velar stop /k/, however, is going to be produced with higher VOT values[17];

b) There are going to be significant differences in both absolute and relative VOT values concerning the quality of the following vowel in the productions of the three places of articulation: differences are going to be not only related to vowel height, but also to vowel frontness[18].

### 3.2.2 L1 American English Production

a) There are going to be significant differences in both absolute and relative VOT values concerning the three places of articulation: the further back the place of articulation, the longer VOT is going to be[19];
b) There are going to be significant differences in both absolute and relative VOT values concerning the quality of the following vowel in the productions of the three places of articulation: differences are going to be not only related to vowel height, but also to vowel frontness;
c) There are going to be significant differences in both absolute and relative VOT values as regards the number of syllables of the target word initiated by the plosive segment: disyllabic words are going to be produced with shorter VOT than monosyllabic ones[20].

### 3.2.3 Brazilian Proficient Learners' Production in English (Interlanguage)

Inasmuch as this study verifies production from Brazilian learners at a proficient level, the hypotheses established for this case were the same three as those for the L1 American English Production. According to what was explained in subsection 2.3, studies have shown that proficient Brazilian learners tend to achieve levels of VOT which are closer to those of

[17] This hypothesis was established based on the findings from Reis & Nobre-Oliveira (2008), Gewer-Borella (2010) and França (2011), discussed in subsection 2.3, which indicate that the velar consonant is produced with longer VOT values in Southern BP.

[18] In this case, since this has not been tested, no hypothesis is established as to which vowel quality causes the VOT of the preceding segment to be longer. This hypothesis is extended to the other two language systems (American English and Interlanguage).

[19] (*cf* LISKER & ABRAMSON, 1964; YAVAS & WILDERMUTH, 2006; YAVAS, 2008)

[20] (*cf* FRANÇA, 2011)

the target language than to the ones found in their mother-tongue. Therefore, it is reasonable to assume that the same effects should be found in the productions obtained from L1 speakers and proficient learners of English.

a) There are going to be significant differences in both absolute and relative VOT values concerning the three places of articulation: the further back the place of articulation, the longer the VOT is going to be;
b) There are going to be significant differences in both absolute and relative VOT values concerning the quality of the following vowel in the productions of the three places of articulation: differences are going to be not only related to vowel height, but also to vowel frontness;
c) There are going to be significant differences in both absolute and relative VOT values as regards the number of syllables of the target word initiated by the plosive segment: disyllabic words are going to be produced with shorter VOT than monosyllabic ones.

### *3.3 Participants*

Two groups of participants integrated the present study. The first group consisted of 10 proficient Brazilian learners of English. All of these were born in the Brazilian southernmost state, Rio Grande do Sul, and had acquired only Brazilian Portuguese before reaching 6 years of age. All of them signed a "*Formulário de Consentimento"* (Appendix 1) and filled out a *"Ficha de Informações do Participante"* (Appendix 2)[21]. The learners (6 male and 4 female) were between 18 and 35 years old, and had not lived in an English-speaking country for more than one and a half years[22]. Three of the ten participants reported to speak

[21] In English, a *Consent Form*, and a *Participant Information Form*, respectively. The consent form was signed by the participants in order to allow the author of this study to use the data provided by them for analysis. By filling out the Participant Information Form, participants provided background information on the language variety they spoke and factors which might influence their production, in order to have a more homogeneous group of participants. Participants from both groups signed and filled out both forms in their native language.

[22] It is understood that participants who lived in an English speaking-country for more than a year may have their production both in L1 and L2 much more influenced by the target language than those who have not had this experience. There was only 1 participant, however, who lived in an English-speaking country for more than 6 months. The mean VOT values verified in this participant's production, when contrasted with those of the other participants, were even lower than the VOT values present in the production of learners who had never left Brazil. Therefore, this is not believed to have played a role which compromises the analysis done by the present study.

other Foreign languages: Spanish (1) in an intermediate level, French (1) and Russian (1) in a basic level – although such proficiency levels were not tested (they were reported by the learners) this is not expected to have affected their VOT values in L2 production, because these languages are not characterized as having aspirated plosive consonants. Their average in terms of age, age at which they started studying English, length of time studying English, and length of residence (LOR) in an English-speaking country are presented in *Table 1* which follows.

Table 1 – Brazilian learners' (n=10) language background information

| Information | Mean | SD |
|---|---|---|
| Age | 22.7 | 5.16 |
| Age at which started learning English | 12 | 3.71 |
| Length of time studying English (years) | 8.5 | 4.5 |
| LOR in an English-speaking country (months) | 2.1 | 5.62 |

All the Brazilian learners were evaluated and classified by the *Oxford Placement Test 1* (ALLAN, 2004)[23]. The ten Brazilian learners who took part the present study were classified in this test as *upper intermediate* (level 6 out of 10), *proficient* (level 7/10) or *highly proficient* (level 8/10) learners of English, and were then considered part of the same group, here labeled as "Proficient". Other participants (two) had been evaluated with a lower proficiency level and were, thus, not selected to take part in this research study.

The second group was formed by 5 native speakers of American English, who had been living in Brazil for no longer than 3 years, and had acquired only American English before reaching 6 years of age.[24] All the participants signed a *Consent Form* (Appendix 3) and

[23] The *Oxford Placement Test 1* (ALLAN, 2004) consists of two parts: the *Grammar Test* and the *Listening Test*. In the *Grammar Test*, participants are assessed concerning their English grammar knowledge through the choice of one correct alternative out of three which could fill in a blank in a text. In the *Listening Test*, learners are evaluated by their capacity to comprehend isolated utterances, choosing one correct alternative out of two, which are key-words in these sentences. For each task a score of 100 points is attributed to the learners' performance, according to their accuracy level. Finally, considering their score, the learners are classified in 1 out of 10 levels, which range from "*beginner*" (level 1, less than 75 points) to "*functionally bilingual*" (level 10, more than 198 points).

[24] There was no control over the variety of American English spoken by these participants, which might be regarded as a limitation of the present study.

responded to a *Participant Information Form* (Appendix 4). The native speakers (4 male and 1 female) were between 20 and 66 years old. Their average age, hours dedicated to the learning of Portuguese weekly, and LOR in Brazil are presented in *Table 2* which follows.

Table 2 – American English native speakers' (n=05) language background information

| Information | Mean | SD |
|---|---|---|
| Age | 36 | 19.53 |
| Hours dedicated to the learning of Portuguese weekly | 11.8 | 10.18 |
| LOR in Brazil (months) | 12.4 | 13.64 |

### *3.4 Selection of target words: English*

A total of 96 English words was selected for the production test: 72 target words (36 monosyllabic and 36 disyllabic) and 24 distractor words (12 monosyllabic and 12 disyllabic). The selection process aimed to obtain the same number of tokens (three) concerning each place of articulation (bilabial, alveolar and velar) and each vowel quality (high-front, low-front, high-back and low-back). A *sound search*[25] was run in the *MacMillan English Dictionary: 2nd Edition* (CD-ROM version) so that a list of all the possible target words could be obtained.

For both monosyllabic and disyllabic English words, the words selected were those initiated by a voiceless stop consonant, preferably following the Consonant + Vowel + Consonant (CVC) syllabic pattern[26]. Vowel + /r/ sequences were avoided, for in this case vowels undergo *r-coloring*[27] and are, thus, more centralized. Moreover, a specific spelling pattern for vowels was sought: for the /u/ vowel, for example, there was an attempt to choose words in which this vowel was represented by the 'oo' spelling. This is justified by the fact that some participants were L2 learners and, inasmuch as English is characterized by a much

[25] This function allows the researcher to look up a word by its pronunciation. A phonetic-alphabet keyboard is presented in order for the researcher to enter the target of the search; once it is performed, the list of all the words existent in the dictionary with that sequence of phonemes is presented.

[26] It was important to keep the same syllabic pattern so that words would have similar length, and vowels would vary less in respect to that aspect. Vowels which are not followed by a final consonant tend to be longer than those which are. Furthermore, consonant clusters in the initial position were avoided, since they could affect VOT. This pattern could not be followed, however, for all of the disyllabic words, since not enough words under that circumstance were found.

[27] As explained by Celce-Murcia *et al* (2010), when followed by that consonant in the same syllable, "the vowel anticipates and glides toward the central /r/ position and takes some of the retroflex quality of /r/" (p. 126).

more opaque grapho-phonic-phonological relationship than Portuguese[28], different spellings could make it more difficult for learners to produce the target vowel in words they were not familiar with.

The target vowels in English were the high-front vowel /i/ (as in "peek"), the high-back vowel /u/ (as in "cool"), the low-front vowel /e/ (as in "pet") and the low-central/back possible vowels /ɑ/ and /ɔ/ (as in "top"). One of the objectives here was to use clearly distinct vowels which would be most similar (that is, close in height and frontness) to those produced in the learners' L1. The high vowels produced by the learners were not expected to diverge much from its target, since, in terms of articulation they can be considered similar: both are produced with similar tongue position and shape. It could be argued that the most adequate target in the low-front vowel position would be the /æ/ vowel, since it is a long vowel, like the high ones. This vowel is, however, most commonly spelled with an 'a', which could lead L2 speakers to pronounce it [eɪ] or [a]. Furthermore, the /e/ and /æ/ targets were expected to be pronounced equally by Brazilian learners, with the Brazilian Portuguese [ɛ] vowel, as in "*pé*" (foot)[29].

It should also be clarified that a high degree of variation can occur in the vowel quality herein labeled as "low-back", both in L1 and L2 productions. In all of the words selected with this vowel quality (presented in the box below), native speakers of American English were expected to vary their production between the vowels [ɑ] and [ɔ]; L2 Brazilian learners would, on the other hand, tend to pronounce it as [ɔ] (similarly to the BP target in that vowel quality), but could as well produce the [ɑ] vowel in a lower frequency. Even though there could be high rates of variation, all of the possible vowels in this case fit a clearly distinct category in comparison to the other target vowels.

As regards disyllabic words, a much more restricted range of possible target words was found. Concerning the case of /p/ followed by a high-back vowel, only two words following the aforementioned desirable conditions were found. The solution was to select a word with the /ʊ/ vowel, *"pudding"*, which should not differ considerably from those with the /u/ vowel, especially because this is a cognate word.

Finally, concerning the distractor words, they were selected with the objective to have as many different initial consonants as possible in the instrument for data collection. Different vowels were also selected in the distractor words, which had no more than two syllables. The

[28] (*cf* ZIMMER, SILVEIRA & ALVES, 2009) See subsection 2.2.

[29] (*cf* SILVA, 2012)

number of distractor words was equivalent to approximately 33% of the number of target words.

In accordance with the criteria expressed above, *Box 1* displays the 72 English target words selected for the present study. The 24 distractor words are displayed in *Box 2*:

Box 1 – English target words (72).

| Consonant | Vowel | Monosyllabic | Disyllabic | Vowel | Monosyllabic | Disyllabic |
|---|---|---|---|---|---|---|
| / p / | high-front | *peace* | *peacock* | high-back | *pooch* | *poodle* |
| | | *peek* | *peanut* | | *poof* | *pootle* |
| | | *peep* | *people* | | *pool* | *pudding* |
| | low-front | *peck* | *pedal* | low-back | *pop* | *pocket* |
| | | *peg* | *pepper* | | *posh* | *polish* |
| | | *pet* | *petal* | | *pot* | *potter* |
| / t / | high-front | *teach* | *teacher* | high-back | *tool* | *toothy* |
| | | *tease* | *teaser* | | *toot* | *tootle* |
| | | *teeth* | *teeny* | | *tooth* | *toucan* |
| | low-front | *tech* | *techno* | low-back | *tod* | *toddle* |
| | | *tell* | *teddy* | | *top* | *topic* |
| | | *ten* | *temper* | | *toss* | *toxic* |
| / k / | high-front | *keel* | *keenly* | high-back | *cool* | *cooler* |
| | | *keen* | *keeper* | | *coop* | *cooper* |
| | | *keep* | *keyword* | | *coot* | *kooky* |
| | low-front | *keg* | *kelvin* | low-back | *cob* | *coffee* |
| | | *ken* | *kennel* | | *cod* | *collar* |
| | | *ketch* | *kettle* | | *cop* | *copper* |

Box 2 – English distractor words (24).

| Monosyllabic | Disyllabic |
|---|---|
| *fight* | *mountain* |
| *house* | *navy* |
| *lap* | *foreign* |
| *moon* | *vessel* |
| *not* | *center* |
| *short* | *zealot* |
| *sock* | *shower* |
| *well* | *winger* |
| *house* | *hunter* |
| *lap* | *lemon* |
| *youth* | *yogurt* |
| *think* | *then* |

### *3.5 Selection of target words: Brazilian Portuguese*

Fourty-four words in Brazilian Portuguese words were selected for the production test, of which 33 were target words and 11 were distractor words. The criteria for determining

which words would be included in the test were similar to the ones used with English words. The aim was to have three tokens for each place of articulation and each target vowel. No particular software or dictionary was used to search for the target words in BP, as this did not prove to be necessary.

Following the same criteria expressed in the previous subsection, four target vowels in Brazilian Portuguese were selected: the high-front vowel /i/, which is spelled with 'i', as in the BP word "*pico*" (*peak)*; the low-front vowel /ɛ/, spelled with 'e', as in "*tela*" (*screen*); the high-back vowel /u/, spelled with 'u', as in "*puro*" (*pure)*; and the low-back vowel /ɔ/, which is spelled with 'o', as in "*cola*" (*glue*). The number of monosyllabic words initiated by a stop consonant and followed by one of these four vowels was not enough to have three tokens of each 'plosive' + 'target vowel' system, considering C+V or C+V+C sequences[30]. Therefore, the production test had only disyllabic BP target words. Every one of the selected target words was composed by two C + V syllables. Only the plosives in the initial syllable had their VOT values measured.

One particular phonological context in BP could not be investigated in this study. As it was mentioned in subsection 3.3, all the Brazilian learners who participated in this study were born in Rio Grande do Sul. In the variety of BP spoken by those participants, whenever the alveolar stops (/t/ and /d/) are followed by the vowel [i], the plosive consonant undergoes *palatalization*[31]: the place of articulation of the vowel is assimilated by the consonant, hence producing an affricate segment (/t/ + [i] = [tʃ]; /d/ + [i] = [dʒ]), as in the BP words "*tipo*" (*type*) and "*disco*" (*disc*). The production of the /t/ + [i] sequence was therefore not tested in BP.

As regards the distractor words, they were chosen according to the same criteria followed in the selection of the English words (explained in the previous subsection). The number of distractor words was equivalent to 33% of the number of target words. *Box 3* displays the 33 BP target words selected for the present study. The 11 distractor words are displayed in *Box 4*:

[30] It is possible that the desired number of target words was acquired if both C+V and C+V+C monosyllabic words were considered together. It was an objective here, however, to maintain the same syllabic pattern, by choosing between one of these syllabic structures (so that there was not much variation in the quality of the target vowels). For that reason, production of monosyllabic words in BP was not tested.

[31] (*cf* OLIVEIRA, 1990; KAMIANECKY, 2002; and MONARETO *et al*, 2005)

Box 3 – BP target words (33).

| Consonant | Vowel | Target Word | Vowel | Target Word |
|---|---|---|---|---|
| **/ p /** | high-front | *pico* | high-back | *puma* |
| | | *pipa* | | *puro* |
| | | *pira* | | *puxe* |
| | low-front | *peça* | low-back | *poça* |
| | | *peço* | | *posso* |
| | | *pega* | | *pote* |
| **/ t /** | high-front | [ tʃ ] | high-back | *tudo* |
| | | | | *tufo* |
| | | | | *tusso* |
| | low-front | *tela* | low-back | *toca* |
| | | *terra* | | *toque* |
| | | *teto* | | *tosa* |
| **/ k /** | high-front | *quibe* | high-back | *cujo* |
| | | *quilo* | | *cume* |
| | | *quina* | | *cura* |
| | low-front | *queda* | low-back | *cola* |
| | | *quepe* | | *copa* |
| | | *quero* | | *copo* |

Box 4 – BP distractor words (11).

| Distractor Word |
|---|
| *choque* |
| *figo* |
| *frasco* |
| *jogo* |
| *lago* |
| *mola* |
| *nota* |
| *rato* |
| *sala* |
| *vaso* |
| *zebra* |

### *3.6 Production test*

The production test consisted of reading aloud the 44 BP words and the 96 English words presented in the previous subsections. BP words were inserted in "*Diga '_____' para mim."* carrier sentences, whereas English words were inserted in "*Say '_____' to me.*" sentences, which are equivalent in meaning to the BP sentences. It is also important to highlight that, in both the English and the BP sentences, the target word was preceded by a vowel and followed by a plosive consonant – thus facilitating the segmentation of the sentence and its subsequent analysis.

Participants were invited to a silent room where the recordings took place. By using a headset[32], participants had their oral reading of the sentences recorded[33] with the software *Audacity* version 2.0.2[34], which had previously been set at stereo quality and at a 22050Hz frequency. The sentences were shown individually through a *Microsoft Power Point* slide presentation.

Participants were instructed to read each one of the sentences aloud, without any pause between its words. They were also instructed to read the sentence again in case they did make a pause between the words, or stuttered, for example. The slides were changed in accordance with the command given by the participants. As regards Brazilian learners' production in English, participants were also instructed to read the sentence again if they changed the syllable stress of the target word.

The order in which the words were presented was previously determined by a randomization, which was performed on *random.org*[35]. After the randomization of the list of target words was performed, the first two distractor words from the resulting list were moved to the first and second positions in the slide presentation. Such a measure had the intention to avoid the loss of tokens due to lack of attention or occasional nervousness of the participants, mostly common in the beginning of the task.

Each word (target or distractor) was read twice, producing a total of 66 tokens of Brazilian Portuguese words per participant; for the English target words, the number of tokens per participant was 144. Sentences with target and distractor words altogether totalized 280 (192 in English and 88 in BP) per participant. They were divided in 6 slide presentations. The first two slide presentations had only BP words; each of these two presentations had 44 sentences divided in two sets, with an optional pause between them. The four other slide presentations had only English words; each of the four presentations had 48 sentences divided in two sets, with an optional pause between them.

---

[32] The headset used in the recordings was *Philips*'s SHM1500. Its specifications are as follows: frequency response: 20 – 20 000 Hz; impedance: 32 Ohm; maximum power input: 100 mW; microphone cartridge: 6mm; sensitivity: 102 dB; speaker diameter: 27 mm; sensitivity microphone: 20-11000Hz, -42+/-3 dB.

[33] The preference for the use of a headset in the recordings is justified by the fact that, with the aid of a headset, the distance between the microphone and the participant's mouths is maintained during the whole recording procedure.

[34] Free software which allows its users to record and edit audio files in various formats. Information can be found at [http://audacity.sourceforge.net/].

[35] This website allows its users to randomize any list of items for free. Information on its operation can be found at [http://www.random.org/].

Brazilian learners of English read all the 6 slide presentations – with the additional time necessary for the participants to take the placement test, the whole procedure lasted approximately 45 minutes for each participant. Native speakers of American English, on the other hand, read only the 4 presentations in English – a procedure of approximately 25 minutes. Brazilian learners produced altogether 880 tokens of BP sentences (660 of which had target words) and 1920 English sentences (1440 of which had target words). Native speakers of American English produced 960 tokens of English sentences (720 of which had target words).

### *3.7 Analysis*

The next stage consisted of the acoustic verification of the audio data in software *Praat* version 5.3.22 (BOERSMA & WEENINK, 2012). Each of the sentences was measured so that two values could be obtained: (i) the absolute Voice Onset Time (VOT) value of the target initial plosive and (ii) the *relative VOT value* (RV): each sentence had its whole length measured and the percentage of time occupied by VOT inside that sentence was calculated to reach this measure; that is, RV = [*VOT length x* 100 / *sentence length*].

The criterion followed in obtaining the VOT value of the initial stops was to measure from the exact beginning of the plosive burst started to the point in which the first pulse of vocal cords could be noticed. *Figure 2* illustrates the measurement of the VOT in the first syllable in the word “kooky” (/k/ + /u/), produced by a native speaker of American English, with 61.80 ms of VOT:

Figure 2 – VOT measurement: /k/ + /u/ sequence with 61.80 ms of VOT.

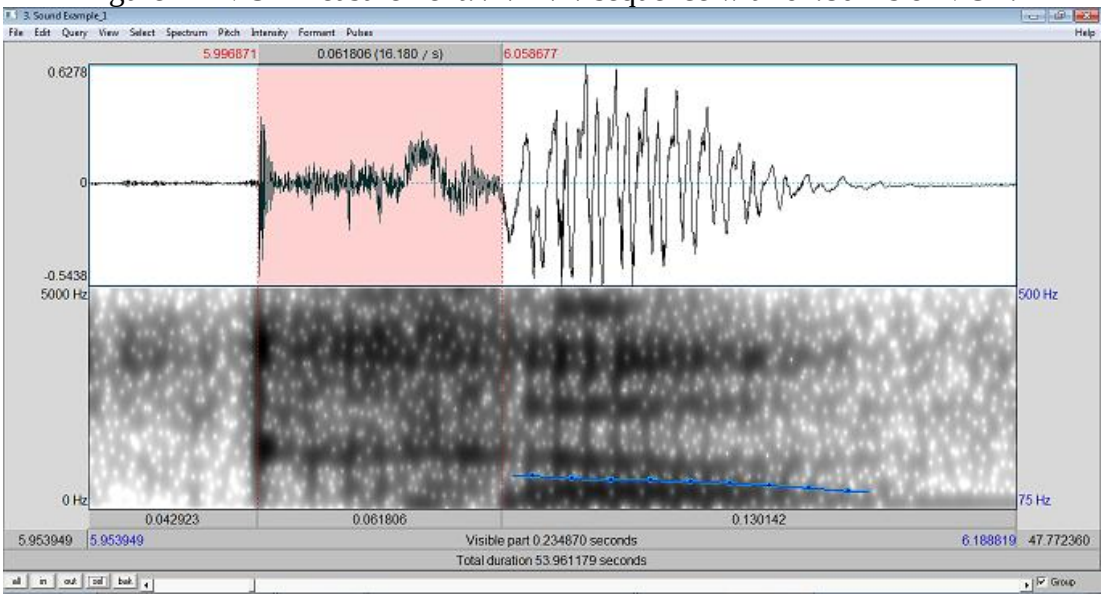
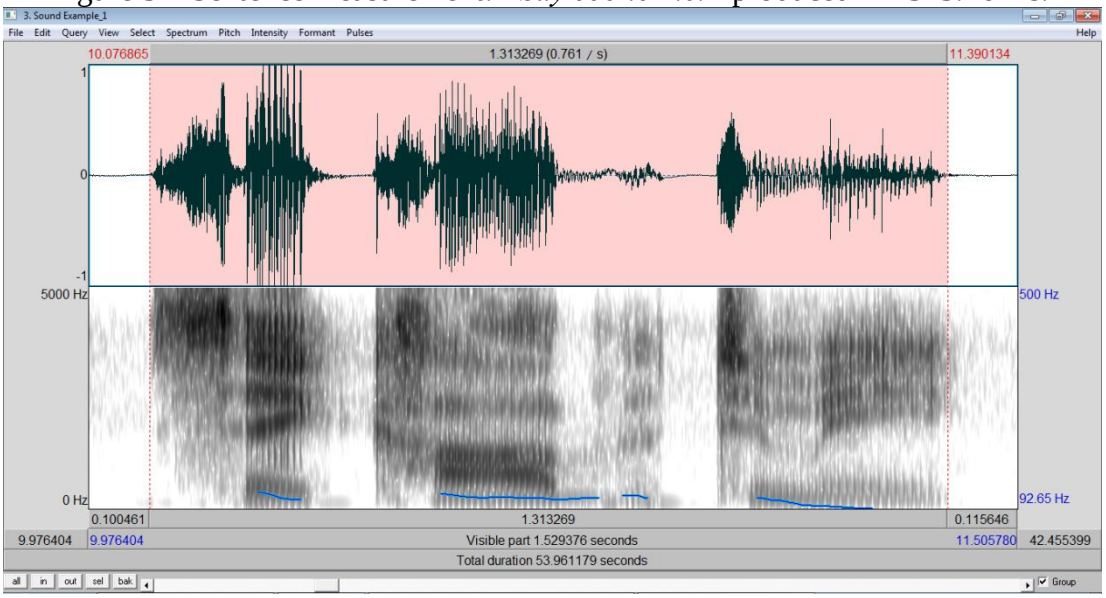


*Figure 3* displays the measurement of the sentence *"Say cob to me."*, produced by a native speaker of American English, with total duration of 1313.26 ms. Sentences in English were measured from the beginning of [s] noise production (in the word "*say"*) to the last glottal pulse of the [i] vowel (in the word *"me"*). Measurement of sentences in BP had an initial point in the first glottal pulse of the *pre-voicing* in [d] production[36] (in the word *"diga"*), and an ending point in the last glottal pulse of the nasalized vowel [iⁿ] (in the word *"mim"*).

Figure 3 – Sentence measurement: *"Say cob to me."* produced in 1313.26 ms.




The VOT and RV values of each token were disposed in *Microsoft Excel* tables which were later input in *IBM SPSS Statistics* version 17.0.0 for statistical analysis. Each of the three cases[37] had a *Multivariate Analysis of Variance* (MANOVA)[38] conducted several times: one for hypothesis (a), concerning place of articulation; three for hypothesis (b), concerning the vowel quality (one per place of articulation); and three others for hypothesis (c), concerning the number of syllables (one per place of articulation). Results for hypotheses (a) and (b) were determined by *post-hoc* Bonferroni's *pairwise comparisons*. In relation to the analysis of

[36] *Pre-voicing* was not detected in a few productions. In that case, the plosive [d] burst was considered as the starting point of the measurement.

[37] The three cases here are those according to which the hypotheses are divided: native Brazilian Portuguese production, L1 American English and Brazilian proficient learners' production in English.

[38] The Multivariate Analysis of Variance is a parametric statistical test used for comparing means of different groups in analysis containing two or more dependent variables (absolute and relative VOT values) and multiple independent variables (in the variable 'place of articulation', for example, we have three possibilities: /p/, /t/ and /k/).

number of syllables and vowel effects, a SPSS feature (*select cases*) was used in order to analyze one place of articulation at a time. The tests displayed, simultaneously, analysis for both measurements absolute and relative VOT values. Results are presented in the next section.

# 4 RESULTS AND DISCUSSION

This section presents the results of the acoustic verification conducted in this study, as well as the statistical analysis carried out to test the hypotheses expressed in subsection 3.2. Results are explained separately for each of the three cases – production in Brazilian Portuguese (L1), American English (L1) and Brazilian English (L2).

It is important to clarify that, for delimitation purposes, the statistical results will only be presented for tests run with the traditional measurement of the absolute VOT value. The statistical results for the tests run with the RV measurement will only be presented in case there is divergence in the results obtained from the two measurements.

Moreover, it is necessary to highlight that not all of the data collected in the production test (quantified in subsection 3.6) could be considered in the analysis. Some tokens were discarded for several different reasons, such as mispronunciation, background noise interference or clear pauses produced between words within a sentence. The total number of tokens considered in each analysis will be expressed in the tables that follow.

## *4.1 Analyzing Brazilian Portuguese Data*

The verified VOT values when analyzing Brazilian Portuguese data are presented, in two measurements, in *Table 3*, which follows:

Table 3 – Mean absolute and relative VOT values in BP production.
VOT: absolute VOT (ms); RV: relative VOT (%); *SD*: standard deviation.

| ***Consonant*** | ***Vowel*** | ***Tokens*** | ***VOT*** | | ***RV*** | |
|---|---|---|---|---|---|---|
| | | | *Mean* | *SD* | *Mean* | *SD* |
| / p / | high-front | 59 | 15.15 | 5.85 | 1.39 | 0.69 |
| | low-front | 59 | 11.57 | 3.85 | 1.02 | 0.41 |
| | high-back | 58 | 19.75 | 6.26 | 1.77 | 0.71 |
| | low-back | 59 | 14.06 | 5.47 | 1.24 | 0.61 |
| / t / | low-front | 59 | 15.32 | 5.01 | 1.35 | 0.50 |
| | high-back | 60 | 23.33 | 7.15 | 2.05 | 0.67 |
| | low-back | 59 | 14.97 | 5.10 | 1.30 | 0.48 |
| / k / | high-front | 55 | 77.22 | 14.80 | 6.72 | 1.17 |
| | low-front | 58 | 46.50 | 12.08 | 3.92 | 1.01 |
| | high-back | 60 | 61.27 | 14.82 | 5.22 | 0.99 |
| | low-back | 60 | 47.23 | 10.60 | 4.06 | 0.78 |

Interesting questions are raised once we observe this table. The VOT values verified for /p/ and /t/ were expected, since they are not distant from those verified in previous studies (REIS & NOBRE-OLIVEIRA, 2008; GEWEHR-BORELLA, 2010; FRANÇA, 2011) – the ones verified in sequences in which those stop segments were followed by a high-back vowel, however, may be considered a little higher than expected. Considering the hypothesis that this vowel quality actually has the effect of causing bilabial and alveolar stops to have longer VOT, this might be a consequence of the fact that the mean VOT value for such a sequence ('plosive' + 'high-back vowel') was never considered in isolation of those with other vowels by previous studies.

Furthermore, the VOT values verified in /k/[39] production, especially those in which this stop segment precedes a high-front vowel, are considerably longer than the ones verified in the abovementioned studies. It is possible, as suggested by Cohen (2004), that the L2 has an influence over this production: since these VOT values correspond to proficient learners' production, it can be supposed that these learners are in a relatively advanced stage of acquisition of *positive VOT* in their English production, and this pattern could, thus, be transferred to their L1 production. Such an explanation becomes more plausible when a factor that facilitates the acquisition of *positive VOT* in this specific place of articulation is taken into account: as addressed in subsection 2.3, VOT values verified in Southern BP for this stop are already higher than those expected of an unaspirated segment.

### *4.1.1 Testing place of articulation effects on VOT in Brazilian Portuguese*

Aiming to test hypothesis (a), presented in subsection 3.2.1, a *MANOVA* test revealed that there are significant differences regarding the place of articulation of the initial stop on the mean VOT values produced in BP: $F$ (2, 87) = 178.486; $p < .05$. Results were also significant with the RV measurement. It is important to clarify that, since there were no tokens of the /t/ + 'high-front vowel' production in BP[40], in order to balance the number of tokens for each place of articulation and to avoid differences in relation to the production of this vowel quality, sequences with the high-front vowel following /p/ or /k/ were not part of the statistical test run here.

[39] Mean /k/ VOT value in BP, considering all vowels (233 tokens): 57,74 ms (*SD*: 18,02).

[40] According to what was explained in subsection 3.5.

*Post-hoc* Bonferroni's *pairwise comparisons* suggested that this significance was only verified in relation to the velar stop (/k/ x /p/ = $p < .05$, /k/ x /t/ = $p < .05$); the /p/ *versus* /t/ contrast was not significant ($p = .674$). The comparison of means with the RV measurement was not divergent.

Such results corroborate hypothesis (a) for BP in subsection 3.2.1: VOT values in BP for the bilabial /p/ and the alveolar /t/ are not significantly different from one another, whereas the velar stop /k/ is produced with significantly longer VOT, regardless of the measurement (with absolute or relative values) used in this verification.

### *4.1.2 Testing vowel quality effects on VOT in Brazilian Portuguese*

*MANOVA* tests were conducted with the objective to verify vowel effects on VOT in the production of each of the three stop consonants /p/, /t/ and /k/ in Brazilian Portuguese. Significant differences concerning vowel quality were confirmed for bilabial [$F$ (3, 36) = 7.464; $p < .05$], alveolar [$F$ (2, 27) = 11.180; $p < .05$] and velar [$F$ (3, 36) = 20.590; $p < .05$] stop consonants. The same results were obtained with the measurement of relative VOT values.

*Post-hoc* Bonferroni's *pairwise comparisons* suggest the existence of significant differences for bilabial and alveolar stops only in the contrasts 'high-back' x 'low-front' and 'high-back' x 'low-back', indicating that the subsequent low-back vowel /u/ causes stops to have considerably longer VOT values. As regards the production of /k/, significant differences in vowel quality can be verified in all but the 'low-front' x 'low-back' vowel quality contrast. Significance values ($p$) in each of the *pairwise comparisons* are displayed in *Table 4* which follows. There was a single divergence (signalized by an asterisk) in the tests run with absolute and relative VOT measurements for the 'high-back' x 'low-back' contrast: with relative VOT values, $p = .187$ for this contrast.

Table 4 – Significance values (*p*) in Bonferroni's *pairwise comparisons* testing vowel quality effects in Brazilian Portuguese production.

| **vowel quality** | **significance value (*p*)** | | |
|---|---|---|---|
| | /p/ | /t/ | /k/ |
| *high-front x low-front* | .375 | - | p < .05 |
| *high-front x high-back* | .064 | - | p < .05 |
| *high-front x low-back* | 1.000 | - | p < .05 |
| *low-front x high-back* | p < .05 | p < .05 | p < .05 |
| *low-front x low-back* | .996 | 1.000 | 1.000 |
| *high-back x low-back* | p < .05* | p < .05 | p < .05 |

These results confirm hypothesis (b) for BP in subsection 3.2.1: there is a suggestion that vowel-frontness (not only vowel-height) related significant differences on VOT means can be verified in BP production. Although the combination /p/ + 'high-back vowel' seems to produce substantially longer VOT than /p/ + 'high-front vowel', the opposite happens when the place of articulation is velar: the /k/ + 'high-front vowel' has much longer VOT than any other plosive + vowel sequence in BP.

Such findings become interesting once we consider that, although the literature (YAVAS & WILDERMUTH, 2006; YAVAS, 2008; FRANÇA, 2011) has traditionally attributed vowel effects on VOT only to height, there is a suggestion that these effects can also be related to frontness. Thus, the necessity for an analysis beyond the dichotomy 'high' *versus* 'low' may be proving valuable. Moreover, the fact that these effects could as well depend on the place of articulation of the stop segment deserves attention: whereas a subsequent high-back vowel causes VOT in bilabial and alveolar stop segments to be longer, it is the high-front vowel which has such an effect in velar plosives. Another question raised by looking at these results lies in the fact that frontness-related effects are only verifiable as regards high vowels: VOT means are considerably similar while comparing low-front and low-back, regardless of the place of articulation. This could be the focus of further investigation.

In the next section, L1 American English production is analyzed in order to verify if the same effects can be observed and, thus, initiate a discussion on the factors influencing VOT investigated by this study.

### *4.2 Analyzing American English Data*

Mean VOT values in the production of the monosyllabic English words by the American participants are shown with two measurements in *Table 5* which follows:

Table 5 – Mean absolute and relative VOT values in American English (L1) production of monosyllabic target words.
VOT: absolute VOT (ms); RV: relative VOT (%); *SD*: standard deviation.

| ***Consonant*** | ***Vowel*** | ***Tokens*** | ***VOT*** | | ***RV*** | |
|---|---|---|---|---|---|---|
| | | | *Mean* | *SD* | *Mean* | *SD* |
| / p / | high-front | 30 | 64.96 | 19.83 | 5.34 | 1.47 |
| | low-front | 26 | 56.58 | 20.71 | 4.76 | 1.44 |
| | high-back | 29 | 70.82 | 22.19 | 5.77 | 1.45 |
| | low-back | 30 | 65.00 | 26.43 | 5.19 | 1.71 |
| / t / | high-front | 29 | 82.24 | 15.14 | 6.68 | 1.22 |
| | low-front | 30 | 77.20 | 19.41 | 6.48 | 1.26 |
| | high-back | 26 | 83.93 | 15.22 | 6.95 | 1.49 |
| | low-back | 30 | 76.16 | 19.31 | 6.26 | 1.24 |
| / k / | high-front | 30 | 99.23 | 15.58 | 8.11 | 1.30 |
| | low-front | 30 | 85.95 | 21.36 | 6.89 | 1.16 |
| | high-back | 25 | 95.45 | 18.32 | 8.08 | 1.68 |
| | low-back | 30 | 86.10 | 20.62 | 6.94 | 1.17 |

It is interesting to notice that, for the bilabial plosive, the higher mean values are those in which the plosive is followed by a back vowel (same as verified in BP), which does not seem to happen regarding the other two places of articulation. Also for /p/ production, it appears that segments followed by the low-front vowel have considerably shorter VOT, in comparison to any other vowel quality. As regards the production of the alveolar and velar stop consonants, mean VOT values seem to fit what is predicted by Yavas & Wildermuth (2006), Yavas (2008) and França (2011): vowel-height effects are clear, and there do not seem to be effects related to vowel frontness. In general, absolute VOT means can be considered slightly higher than those presented by the literature.

In relation to relative VOT values, the means verified here are similar to those found by Prestes (2012) only in /t/ productions, for which the author found a mean 6,31% relative

value[41]. As regards /p/ and /k/ production (when followed by high vowels, especially), however, these findings show higher relative VOT values than the ones in Prestes (*op cit*), who verified a mean 4,69% RV for /p/ and a mean 6,62% for productions of /k/. The target sentence used in that author's study, it is important to highlight, was the same than the one from the present study.

The next Table (6) presents VOT values in the production of the disyllabic target words by the North-American participants.

Table 6 – Mean absolute and relative VOT values in American English (L1) production of disyllabic target words.
VOT: absolute VOT (ms); RV: relative VOT (%); *SD*: standard deviation.

| ***Consonant*** | ***Vowel*** | ***Tokens*** | ***VOT*** | | ***RV*** | |
|---|---|---|---|---|---|---|
| | | | *Mean* | *SD* | *Mean* | *SD* |
| / p / | high-front | 30 | 64.53 | 17.13 | 5.08 | 1.45 |
| | low-front | 30 | 55.11 | 21.14 | 4.44 | 1.32 |
| | high-back | 28 | 65.90 | 20.92 | 5.25 | 1.23 |
| | low-back | 26 | 64.09 | 23.18 | 4.84 | 1.35 |
| / t / | high-front | 28 | 82.03 | 16.12 | 6.46 | 1.30 |
| | low-front | 30 | 68.23 | 15.19 | 5.57 | 1.30 |
| | high-back | 29 | 78.51 | 16.47 | 6.22 | 1.38 |
| | low-back | 29 | 81.05 | 29.24 | 6.30 | 1.96 |
| / k / | high-front | 27 | 94.36 | 13.84 | 7.35 | 1.38 |
| | low-front | 30 | 86.70 | 27.42 | 6.70 | 1.38 |
| | high-back | 28 | 88.19 | 24.29 | 7.11 | 1.34 |
| | low-back | 30 | 86.85 | 21.10 | 6.89 | 1.27 |

By analyzing *Table 6*, we see that the effects related to vowel height in disyllabic words are not as clear as they were in monosyllabic words production: in fact, only one vowel quality is highlighted as deviant in the production of each of the stop consonants – the low-front vowel for the bilabial and alveolar plosives (with considerably shorter VOT) and the high-front vowel for the latter one (with longer VOT). The higher VOT values for /k/ when followed by a high-front vowel was an effect already verified in the production of monosyllabic words both in Brazilian Portuguese and American English production. The fact that the VOT values observed in the production of this sequence differ from the ones verified

[41] Although the author did not verify vowel height effects, the target words in this study (all monosyllabic) considered three vowel qualities: high-front, low-front and high-back.

in /k/ + 'high-back vowel' production may suggest, once again, the existence of vowel-frontness related effects on VOT.

Once we verified similarities in production of both BP and American English, some explanations can be offered. First of all, it is necessary to reassure what was discussed in subsection 2.5: front and back vowels differ considerably in terms of articulation. It could be seen that, in the production of velar stops in both BP and American English, VOT values are higher when /k/ follows a high-front vowel. However, the opposite happens with the bilabial place of articulation: higher VOT values are verified, both in BP and American English, when bilabial /p/ stops are followed by a high-back vowel. It is possible that these results are related to the articulation of these sequences.

The place of articulation in the production of a high-back vowel is clearly close to the velar place of articulation, whereas there is a distance between the place of articulation of a high-front vowel production and /k/ place of articulation. As already discussed (subsection 2.4), the slower articulators undo the obstruction of the oral cavity, the lower the degree of abruptness of the pressure drop in the plosive production is going to be and, consequently, the longer VOT values are going to be. Since the movement made by the tongue in the production of the /k/ + 'high-back vowel' is minimum in comparison to /k/ + 'high-front vowel', there is a higher degree of abruptness in the pressure drop when that sequence is produced – we could, therefore, expect VOT to be shorter in the /k/ + 'high-back vowel' sequence.

The same explanation could be offered for the production of bilabial consonants. The exact opposite to what was described in the previous paragraph might occur, since the position of the tongue in the articulation of high-back vowels is considerably distant from the bilabial place of articulation.

The difference in mean VOT values show, however, that the aforementioned effects (high-back vowels causing VOT to be longer in /p/, but /k/ showing higher VOT values when followed by high-front vowels) are more clear when we analyze BP data than American English Data. Cho & Ladefoged (1999) alluded to the fact that some explanations to the variation in VOT values are more satisfactory for languages with unaspirated stops. Among these are the explanation related to the volume of the cavity in front of the point of constriction and the one regarding the movement of articulators[42]. It might be argued, therefore, that the vowel-quality related effects verified here are clearer in BP production because this language system is characterized by an unaspirated voicing pattern – and thus, it

[42] Postulates (2) and (3) out of 6 explanations offered by the authors for variation in VOT values. Further explanation can be found in Cho & Ladefoged (1999).

is possible to explain VOT variation by, for example, comparing the movements of the articulators in the CV sequences, as we attempted in the previous paragraphs. Since American English stop consonants are aspirated, these explanations may not be valid. Deeper acoustic analysis seems to be necessary in this sense.

As regards the alveolar place of articulation, the articulatory explanations discussed here do not seem valid, for mean VOT values are very similar in the production of the sequences involving high-front and high-back vowels in English – since there was no /t/ + /i/ production in BP, it is important to look at the interlanguage data, on which this sequence may be found, and establish comparisons in this sense. The long VOT in the /t/ + 'low-back vowel' sequence, especially for disyllabic words, deserves attention, for there is an extensive horizontal movement done by the tongue in the articulation of this sequence.

Concerning relative VOT values, the same verified for monosyllabic words could be observed in disyllabic word production: in general, the means verified are higher than those found by Prestes (2012). They are, however, more similar, but the fact that Prestes (*op cit*) only measured sentences with monosyllabic target words must be highlighted, for this is crucial for interpreting relative values: since the sentences have an extra syllable, the relative value is expected to be lower in disyllabic target words.

### *4.2.1 Testing place of articulation effects on VOT in American English*

With the objective to test hypothesis (a) for American English (L1), presented in subsection 3.2.2, a *MANOVA* test revealed that there are significant differences regarding the place of articulation of the initial stop on the mean VOT produced by native speakers of American English in their L1: $F$ (2, 117) = 26,163; p < ,05. Results were also significant with the RV measurement. This test, it is important to highlight, aggregated productions with both monosyllabic and disyllabic target words.

*Post-hoc* Bonferroni's *pairwise comparisons* suggested that all the contrasts established between each of the three places of articulation are significant (/p/ x /t/ = p < .05; /p/ x /k/ = p < .05; /t/ x /k/ = p < .05). The comparison of means with the RV measurement was not divergent.

Such results corroborate hypothesis (a) in subsection 3.2.2: there are significant differences in VOT and RV between the three places of articulation: the further back the place of articulation, the longer the VOT length is. These results are in accordance with what was stated by Lisker & Abrahamson (1964).

### *4.2.2 Testing vowel quality effects on VOT in American English*

*MANOVA* tests were conducted aiming to verify vowel effects on VOT values in the production of each of the three stop consonants /p/, /t/ and /k/ in American English (L1). According to the results obtained, no significant differences concerning vowel quality were found in VOT levels for bilabial [$F$ (3, 36) = .627; p = .602], alveolar [$F$ (3, 36) = .916; p = .443] or velar [$F$ (3, 36) = .895; p = .453] stop consonants. The results obtained with the measurement of relative VOT values were also not significant. There was, therefore, no need to run *post-hoc* Bonferroni's *pairwise comparisons*, as the former test had indicated that no significant differences would be found in any vowel-quality contrast established.

Hypothesis (b) for L1 American English (subsection 3.2.2) was therefore not confirmed: regardless of the measurement used in the analysis, there are no significant differences concerning the vowel quality in productions from any of the three places of articulation. Although Yavas & Wildermuth (2006) and Yavas (2008) did not test L1 production of English, this is opposite to what the authors hypothesized, since, according to them, differences should be significant, due to vowel-height effects.

One can observe that the means displayed in Tables 5 and 6, in subsection 4.2, suggest the existence of significant differences, especially contrasting VOT means for /p/ followed by high-back (mean VOT 70.82 ms) and low-front (mean VOT 56.58 ms) vowels in monosyllabic words, for example. It is likely that the test did not consider these mean differences to be significant due to the small number of participants (5 in total); it is reasonable to expect, therefore, that, with a larger number of participants, such differences would be considered significant. It could also be argued that the differences were not significant because monosyllabic and disyllabic words were considered together in this analysis. Nonetheless, the results presented in the subsection that follows might refute this possibility.

### *4.2.3 Testing the effects of the number of syllables of the target word on VOT in American English*

*MANOVA* tests were conducted for verifying the existence of significant effects in relation to the number of syllables of the target words on VOT in the production of American English. No significant results were found contrasting the two variables (monosyllabic and disyllabic) for bilabial [$F$ (1, 38) = .186; p = .669], alveolar [$F$ (1, 38) = .216; p = .645] or

velar [$F$ (1, 38) = .505; p = .482] stop consonants. Results for the relative VOT values showed no discrepancy.

Hypothesis (c) for American English (L1), expressed in subsection 3.2.2, was thus not corroborated: there are no significant differences in VOT or RV as regards the number of syllables of the target word. França (2011) – while testing Brazilian leaners' production in English, not native production, as tested here – ran a different statistical test (the non-parametric *Wilcoxon* test); a significant difference in mean VOT values was verified, but only for productions of the alveolar stop consonant. It is possible that, in American English production (L1), the same effects cannot be observed. Since this has never been tested in previous studies, it seems like further investigation in this specific point is necessary before concrete conclusions can be made.

### *4.3 Analyzing Interlanguage Data*

Mean VOT values concerning the production of monosyllabic English target words by the proficient learners of English who took part in this study are shown with two measurements (absolute and relative VOT values) in *Table 7* below.

Table 7 – Mean absolute and relative VOT values in Brazilian English (L2) production of monosyllabic target words.
VOT: absolute VOT (ms); RV: relative VOT (%); *SD*: standard deviation.

| ***Consonant*** | ***Vowel*** | ***Tokens*** | ***VOT*** | | ***RV*** | |
|---|---|---|---|---|---|---|
| | | | *Mean* | *SD* | *Mean* | *SD* |
| / p / | high-front | 59 | 21.30 | 14.84 | 1.97 | 1.26 |
| | low-front | 58 | 16.72 | 10.57 | 1.55 | 0.95 |
| | high-back | 58 | 33.12 | 21.17 | 3.05 | 2.00 |
| | low-back | 58 | 22.49 | 11.94 | 2.08 | 1.07 |
| / t / | high-front | 58 | 43.72 | 21.39 | 3.89 | 1.89 |
| | low-front | 58 | 46.27 | 21.09 | 4.40 | 2.13 |
| | high-back | 55 | 47.90 | 21.03 | 5.05 | 6.41 |
| | low-back | 60 | 41.12 | 18.36 | 3.86 | 1.83 |
| / k / | high-front | 54 | 85.12 | 18.89 | 7.96 | 2.03 |
| | low-front | 57 | 60.06 | 15.17 | 5.59 | 1.60 |
| | high-back | 55 | 70.66 | 18.46 | 6.52 | 1.68 |
| | low-back | 59 | 57.51 | 15.08 | 5.45 | 1.41 |

Mean VOT values in the proficient Brazilian learners' production of disyllabic English target words is shown with two measurements in *Table 8* which follows.

Table 8 – Mean absolute and relative VOT values in Brazilian English (L2) production of disyllabic target words.
VOT: absolute VOT (ms); RV: relative VOT (%); *SD*: standard deviation.

| ***Consonant*** | ***Vowel*** | ***Tokens*** | ***VOT*** | | ***RV*** | |
|---|---|---|---|---|---|---|
| | | | *Mean* | *SD* | *Mean* | *SD* |
| / p / | high-front | 60 | 24.75 | 12.34 | 2.09 | 0.93 |
| | low-front | 56 | 19.13 | 10.71 | 1.69 | 0.89 |
| | high-back | 50 | 34.45 | 20.52 | 3.04 | 1.74 |
| | low-back | 55 | 26.41 | 11.96 | 2.37 | 1.10 |
| / t / | high-front | 58 | 48.80 | 20.16 | 4.32 | 1.70 |
| | low-front | 58 | 38.48 | 17.91 | 3.32 | 1.52 |
| | high-back | 54 | 44.30 | 22.37 | 3.84 | 2.01 |
| | low-back | 54 | 40.41 | 19.41 | 3.97 | 1.83 |
| / k / | high-front | 59 | 77.36 | 17.86 | 6.55 | 1.78 |
| | low-front | 54 | 62.44 | 15.50 | 5.35 | 1.46 |
| | high-back | 57 | 62.00 | 17.66 | 5.45 | 1.42 |
| | low-back | 49 | 63.25 | 17.61 | 7.99 | 1.69 |

Although an investigation on the acquisition of *positive VOT* by Brazilian learners was not the primary focus in this study, it is important to highlight that VOT means in their production are clearly higher in English (L2) than in BP – indicating that these learners achieved a stage of, at least, *partial* acquisition of the L2 voicing pattern. More importantly, as confirmed by the statistical test presented in the next subsection, learners have succeeded to go from a linguistic system in which there is no distinction between /p/ and /t/ as regards VOT (BP), to a linguistic system in which there is a clear distinction as to VOT values in the production of those consonants. Moreover, the high VOT values in production of /k/ in BP seem to have facilitated native-like aspiration of that consonant, especially for the /k/ + 'high-front vowel' sequence, in which the verified VOT values in L2 were very close to the 80 ms taken as the standard of the target language (*cf* LISKER & ABRAHAMSON, 1964).

França (2011) findings in production of advanced Brazilian learners of English (/p/ = 36,07; /t/ = 56,28; /k/ = 91,71)[43] show similar, though slightly higher VOT values than those

[43] Mean VOT values verified considering all subsequent vowels altogether.

verified in the present study. It should be clear that the group of advanced learners in França (2011) consisted of four learners (in a total of 20), which indicates that França's mean VOT values correspond to a considerably smaller number of tokens.

It can also be highlighted that the vowel quality which caused plosives to show longer VOT values in the learner's L1 production of both monosyllabic and disyllabic target words were the same as in their L2: high-back vowel for /p/ and /t/, but high-front vowel for /k/. This reinforces the idea discussed in the previous subsections, according to which such results might be related to the articulation of these CV sequences. In this sense, it is interesting to notice that these effects were much clearer in the productions by Brazilian learners (in L1 and L2). Resuming the discussion conducted in subsection 4.2, based on what was postulated by Cho & Ladefoged (1999), we might suggest that this is related to the fact that BP stops are unaspirated. Interestingly, it seems like the Brazilian learners might have transferred the vowel-quality variation found in their L1 (high-back vowels causing VOT to be longer in /p/, but /k/ showing higher VOT values when followed by high-front vowels) to their L2 production. However, further investigation is required in this sense, with more data from native speakers of American English.

Another fact that deserves attention is that very little variation was verified between the production of monosyllabic and disyllabic target words. The only considerable deviation was verified between the production of the /k/ + 'high-front vowel' sequence, in which monosyllabic words presented longer VOT – an effect predicted by França (2011), but not consonant with what had been suggested by the findings in Yavas & Widermuth (2006) and Yavas (2008).

As regards the relative VOT values, the present findings indicate considerably higher RV than the ones verified by Prestes (2012) in Brazilian learners' production in English, comparing either monosyllabic or disyllabic target words. Prestes (*op cit*) found mean relative VOT values of 0,74% for /p/, 1,85% for /t/ and 3,75% for /k/ production. A justification for such a divergence may lie in the fact that Prestes (2012) verified production by pre-intermediate learners, whereas the present study investigated production by proficient learners.

Finally, it must be clarified that the lexical frequency of the target words in English was not verified, due to a limitation concerning the time to conduct the present study. This might be regarded as a limitation of this study, especially once we observe high levels of standard deviation on the mean VOT values in Brazilian production in English. Frequent words (that is, words which we expect L2 learners to be familiar with) might be pronounced

more spontaneously than infrequent ones, which might have an effect on the results verified here.

### *4.3.1 Testing place of articulation effects on VOT in interlanguage data*

A *MANOVA* test was conducted with the objective of testing hypothesis (a) for Brazilian learners' production in English (subsection 3.2.3). The statistical test confirmed that there are significant differences regarding the place of articulation of the initial stop on the mean VOT values produced by Brazilian learners in English: $F$ (2, 237) = 157.341; $p < .05$. There was no discrepancy in relation to measurement: results were also significant with the relative VOT values measurement.

*Post-hoc* Bonferroni's *pairwise comparisons* suggested that all the contrasts established between each of the three places of articulation are significant (/p/ x /t/ = $p < .05$; /p/ x /k/ = $p < .05$; /t/ x /k/ = $p < .05$). The statistical analysis with the RV measurement was not divergent. The distinction between /p/ and /t/ is not verified in the participant's production in their mother tongue (see subsection 4.1.1), yet it is observed in the production of native speakers in their target language (subsection 4.2.1). Therefore, it is interesting to notice that, as already addressed in the previous subsection, these results suggest that the Brazilian learners who participated in this study succeeded in acquiring the distinction existent in English in terms of VOT between the production of bilabial and alveolar consonants.

These results corroborate hypothesis (a) in subsection 3.2.3: there are significant differences in absolute and relative VOT values considering the three places of articulation: the further back the place of articulation, the longer the VOT length is.

### *4.3.2 Testing vowel-quality effects on VOT in interlanguage data*

In order to test hypothesis (b) for Brazilian learners' production in English (subsection 3.2.3), *MANOVA* tests were conducted with the objective to verify vowel effects on VOT in Brazilian learners' production of each of the three stop consonants /p/, /t/ and /k/ in English. Results suggest that significant differences concerning vowel quality exist in VOT levels for bilabial [$F$ (3, 76) = 5.316; $p < .05$] and velar [$F$ (3, 76) = 9.737; $p < .05$] stop consonants. As regards alveolar stop consonants, differences were not significant: $F$ (3, 76) = .664; $p = .577$. Results for the analysis with RV were divergent only in relation to velar plosives, whose differences were not significant [$F$ (3, 76) = 1.043; $p = .378$].

*Post-hoc* Bonferroni's *pairwise comparisons* (*Table 9)* indicated significant differences only for the 'low-front vowel' x 'high-back vowel' contrast in /p/ production; the /p/ + 'high-back vowel' sequence has, therefore, the longest VOT for that consonant in Brazilian learners' English. There was a discrepancy in the 'high-front' x 'high-back' vowel quality contrast in relation to the measurement used in the analysis: with the RV measurement, that contrast was significant ($p < .05$). As to the velar consonant production, the /k/ + 'high-front vowel' sequence has substantially longer VOT in contrast with any other vowel quality; other contrasts were not significant.

Table 9 – Significance values (*p*) in Bonferroni's *pairwise comparisons* testing vowel quality effects in Brazilian learners' production in English.

| **vowel quality** | **significance value (*p*)** | |
|---|---|---|
| | /p/ | /k/ |
| *high-front* x *low-front* | 1.000 | p < .05 |
| *high-front* x *high-back* | .067* | p < .05 |
| *high-front* x *low-back* | 1.000 | p < .05 |
| *low-front* x *high-back* | p < .05 | 1.000 |
| *low-front* x *low-back* | .717 | 1.000 |
| *high-back* x *low-back* | .130 | 1.000 |

These results do not confirm hypothesis (b) for Brazilian learner's English (subsection 3.2.2): although there are significant differences in absolute and relative VOT values concerning the vowel quality in stop consonants produced in English by Brazilian learners, these differences are only verified in productions of bilabial and velar stop consonants. Nevertheless, it is important to highlight that the significances are not only related to vowel-height, but also to vowel-frontness: productions of /k/ + 'high-front' are significantly different from those of /k/ + 'high-back' vowels.

The findings presented in this subsection are opposite to what was verified by Yavas & Wildermuth (2006), Yavas (2008) and França (2011). According to these authors' findings, differences on the effects of the following vowel on the VOT of stop consonants would be verified in all of the three places of articulation; moreover, these effects would be related only to vowel height. This once again supports the idea that was discussed in the previous subsections: in all of the three language systems investigated in this study (although not as clearly in American English), the differences in the articulation of high-front and high-back

vowels seem to affect the VOT of the preceding consonant, and these effects are not equally observed in all of the three places of articulation.

### *4.3.3 Testing the effects of the number of syllables of the target word on VOT in interlanguage data*

*MANOVA* tests were conducted for verifying the existence of significant effects in relation to the number of syllables of the target words on VOT in Brazilian learners' production in English. No significant results were found contrasting the two variables (monosyllabic and disyllabic) for bilabial [$F$ (1, 78) = .594; p = .443], alveolar [$F$ (1, 78) = .197; p = .685] or velar [$F$ (1, 78) = .154; p = .695] stop consonants. Results with the relative VOT values were not significant for the productions of the three consonants either.

Hypothesis (c) for Brazilian learners' production in English, expressed in subsection 3.2.3, was thus not confirmed: there are no significant differences in VOT or RV as regards the number of syllables of the target word. As already discussed in subsection 4.2.3, França ran a different statistical test and verified significant differences in mean VOT values only for productions of the alveolar stop consonant. Whereas previous studies (YAVAS & WILDERMUTH, 2006; YAVAS, 2008; and FRANÇA, 2011) suggest that there is a difference in mean VOT as regards the number of syllables of the target word, the present study did not verify such an effect.

## *4.4 Summary of findings in the present study*

Before this section is finalized and a conclusion is reached, it is important to briefly summarize what was found in the present study. Firstly, as regards place of articulation effects on VOT, it could be seen that, whereas American English has clearly different VOT for each of the three places of articulation, BP does not distinguish between bilabial and alveolar stop segments in terms of VOT. By analyzing the interlanguage data of Brazilian learners' production in English, we could verify that these learners managed to acquire the distinction in terms of VOT between the bilabial and alveolar stop consonants – showing, thus, significantly different VOT values for each of the three places of articulation.

Once we look at the results concerning the effects of the quality of the subsequent vowel on VOT, we can observe that, in general, there is a distinction between the VOT values of consonants followed by high-front and high-back vowels – which might be justified by the

differences in their articulations. Moreover, vowel quality effects on VOT do not seem to be equal in all of the three places of articulation: whereas /k/ + 'high-front vowel' sequences have the longest VOT in velar plosive production, for bilabial plosives, the higher VOT values are verified when these segments are followed by high-back vowels. This effect can be verified in the three language systems investigated here, although they are not very clear in American English production (L1).

In relation to the effects of the number of syllables of the target word on VOT, it could be seen that there are no significant differences concerning the VOT values found in the production of monosyllabic and disyllabic words in both American English and Brazilian English.

Finally, regarding the possible divergences between the analyses conducted with absolute VOT values and the ones ran with relative VOT values, it can be stated that very few cases of discrepancy were verified.

## 5 CONCLUSION

The present study investigated the production of plosive stops and three factors influencing Voice Onset Time (namely place of articulation, vowel quality and number of syllables of the target word) in three language systems: Brazilian Portuguese, American English and Proficient Brazilian learners' English. Production data were measured using absolute and relative VOT values in order to verify divergences in the results obtained with the use of these measurements.

In general, the mean VOT values verified here in BP are in accordance with those found in the literature (REIS & NOBRE-OLIVEIRA, 2008; GEWEHR-BORELLA, 2010; FRANÇA, 2011) – except for the ones verified in velar plosive production, which are higher than those from Reis & Nobre-Oliveira (2008) and França (2011). The mean VOT values verified in American English (L1) production are also higher than those predicted by the literature (LISKER & ABRAHAMSON, 1964). Brazilian learners' production in English shows mean VOT values which are considerably higher than those found in their L1 – especially in /k/ production –, indicating that these learners had acquired, at least partially, the *positive VOT* of target language.

The mean relative VOT values observed in this study were higher than those verified by Prestes (2012). Furthermore, very few discrepancies were found between the results obtained from the tests run with absolute VOT values and those run with relative VOT values. This suggests that it may be valuable for further investigations on this phonetic-phonological aspect to use both measurements, since the analysis with relative VOT values (not traditionally adopted) might account for speech-rate related problems on measuring VOT.

As regards place of articulation effects on VOT, our analyses show that in BP no distinction is made in terms of VOT values between the bilabial and alveolar stop segments, whereas the VOT found in these two segments differ considerably from the one verified in the velar plosive production. In American English, however, there was a clear distinction between the VOT values found in the productions of each of the three places of articulation. Analysis on the production in English by proficient Brazilian learners show that participants managed to acquire the distinction on VOT values between bilabial and alveolar stops, in a similar fashion to the distinction observed in the target language, thus producing clearly distinct VOT levels in the three places of articulation.

Concerning vowel quality effects on VOT, the analyses conducted in this study show something which had not yet been verified by the literature (YAVAS & WILDERMUTH,

2006; YAVAS, 2008; FRANÇA, 2011), as it is shown that the effects of the subsequent vowel on VOT might not only depend on vowel height, but also on vowel frontness. Moreover, these effects are not the same in the three places of articulation: while a subsequent high-back vowel seems to favor longer VOT in bilabial stops, it is the high-front vowel which causes velar plosive segments to show longer VOT. Although this is clear in the L1 and L2 production by Brazilian learners of English, the same cannot be stated concerning the L1 American English production. Also, it could be argued, there is a possibility that these effects be explained by the differences in the co-articulation between consonant and vowel in these CV sequences. In this sense, further investigation seems to be required, aiming to explain the reason why frontness plays a role on VOT distinction between high vowels, but not between low vowels.

The suggestions made in this study regarding vowel quality effects could not always find justification in statistical analysis, however: statistical analysis confirms effects related to vowel quality on VOT in BP and Brazilian English production, but not in L1 American English production, even though considerable differences were observed in the mean VOT values – once the analysis considering vowel height and frontness separately might be proving valuable, future studies should test these effects in a larger sample, especially in American English. The explanation offered here, that these effects might be clearer in language systems with unaspirated segments – based on what was postulated by Cho & Ladefoged (1999) – should also be taken into account.

As to the effects of the number of syllables of the target word on VOT, no significant differences on the mean VOT values in the production of monosyllabic and disyllabic words could be verified in any of the three language systems involved. This is not consonant with what had been suggested by previous studies, that monosyllabic words tend to have longer VOT (FRANÇA, 2011) and, in opposition, that disyllabic words might have longer VOT (YAVAS & WILDERMUTH, 2006).

A limitation of the present study can be found in the small number of participants in the group of native speakers of American English, as well as in the little control over the variety of American English spoken by these participants, since they were born in different regions of the United States, which might influence the VOT levels found in their production. This leaves a gap for future studies: further investigation could be made with a larger group of participants, speakers of the same variety of American English.

Another factor which might be regarded as a limitation of the present study is the fact that the lexical frequencies of the English target words have not been verified. As already

discussed in this text, especially because high levels of standard deviation on the mean VOT values are observed in the production in English by Brazilian learners, there is a possibility that the frequency of the target word influences the way L2 learners pronounce it.

Future studies may also find an interesting object of study in the influence exerted by the learners' L2 in their L1 production as regards VOT. As addressed in the previous section, the fact that these proficient learners of English produce stop segments with considerably long VOT in their L1 might be might be justified by the fact that they have partially acquired the voicing pattern of their L2 and are transferring this pattern to their L1 production – a possibility raised by Cohen (2004).

Finally, it can be concluded that the present research study has reached its goal of deepening the understanding on the factors influencing Voice Onset Time, since this study conducted analysis and raised questions which had never been considered before by the literature. We expect that these findings motivate further investigations on VOT, the factors influencing this phonetic-phonological aspect and the way it is presented in different language systems.

## REFERENCES

ALBANO, Eleanora Cavalcante. *O gesto e suas bordas: esboço da fonologia acústico-articulatória do português brasileiro.* Campinas: Mercado de Letras, ALB, FAPESP, 2001.

_____. *A pulsação sobre a letra: pela quebra de um silêncio histórico no estudo do som de fala*. Cadernos de Estudos Linguísticos, v.1, n.42, p. 7-20, 2002.

ALVES, Mariane Antero. *Production of English and Portuguese Voiceless stops by Brazilian EFL Speakers*. Dissertação (Mestrado em Letras) - Universidade Federal de Santa Catarina, Florianópolis, SC, 2011.

ALVES, Ubiratã Kickhöfel. Uma discussão conexionista sobre a explicitação de aspectos fonético-fonológicos da L2: dados de percepção e produção da plosiva labial aspirada do inglês. *In*: POERSCH, José Marcelino; ROSSA, Adriana Angelim (org.). *Processamento da Linguagem e Conexionismo.* Santa Cruz do Sul: Editora da UNISC, 2007, p. 155-185.

_____. *Efeitos da Instrução Formal na Aquisição de Aspectos Fonético Fonológicos do Inglês (L2) por Brasileiros***.** Projeto de pesquisa. Porto Alegre: Universidade Federal do Rio Grande do Sul (UFRGS), 2010.

_____; MAGRO, Vivian. Raising awareness of L2 phonology: explicit instruction and the acquisition of aspirated /p/ by Brazilian Portuguese speakers. *Letras de Hoje*, Porto Alegre, v. 43, n. 6, 2011.

_____; SCHWARTZHAUPT, Bruno M; BARATZ, Ana Hemmons. *Percepção e Produção dos Padrões de VOT do Inglês (L2) por Aprendizes Brasileiros.* Estudos em Aquisição Fonológica, vol. 4. Pelotas: Editora Universitária/UFPEL, 2011.

ASHBY, Michael; MAIDMENT, John. *Introducing Phonetic Science.* Cambridge: Cambridge University Press, 2005.

BEST, C; TYLER, M. D. *Nonnative and second-language speech perception: commonalities and complementarities.* In: BOHN, O. S.; MUNRO, M. J. *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2007, p. 13-34.

CELCE-MURCIA, Marianne; BRINTON, Donna M.; GOODWIN, Janet M; GRINER, Barry. *Teaching pronunciation: a coursebook and reference guide*. Cambridge University Press, 2010.

CHO, Taehong; LADEFOGED, Peter. Variation and Universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 1999, p. 207-229.

COHEN, Gustavo Vargas. *The VOT Dimension: a bi-directional experiment with English Brazilian Portuguese stops*. Dissertação – Mestrado em Letras. UFSC, Florianópolis: 2004.

DE BOT, Kes; LOWIE, Wander; VERSPOOR, Marjolijn. *A Dynamic Systems Theory approach to second language acquisition.* Bilingualism: Language and Cognition, v. 10, n. 1, p. 7-21, 2007.

FLEGE, J. E; MUNRO, M. J.; MacKAY, I. R. A. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, v. 97, n.5, p. 3125-3134, 1995.*

FRANÇA, Karoline. *A aquisição da aspiração das plosivas surdas do inglês por falantes do português brasileiro: implicações teóricas decorrentes de duas diferentes formas de descrição dos dados*. Dissertação (Mestrado em Letras) - Universidade Católica de Pelotas, Pelotas, RS, 2011.

KAMIANECKY, Fernanda. *A palatalização das oclusivas dentais /t/ e /d/ nas comunidades de Porto Alegre e Florianópolis: uma análise quantitativa.* Dissertação: Mestrado em Letras, PUCRS. 114 f. Porto Alegre, 2002.

KENT, Raymond D; READ, Charles. *The acoustic analysis of speech*. Los Angeles, CA: Singular/Thomson Learning, v. 2, 2002.

LADEFOGED, Peter. *Vowels and Consonants: An Introduction to the Sounds of Languages.* Los Angeles, CA: Blackwell Publishing, v. 2, 2005.

LISKER L; ABRAMSON, A. *A Cross-language study of voicing in initial stops: Acoustical measurements.* Word, New York, United States, v.3, n.20, p. 384-422, 1964.

MAGRO, Vivian. *Effects of different types of instruction on the acquisition of aspiration in word-initial /p/ by Brazilian Portuguese speakers.* Monografia de conclusão (Especialização em Ensino de Língua Inglesa) - Centro Universitário Ritter dos Reis, Porto Alegre, RS, 2010.

MONARETTO, Valéria N O; QUEDNAU, Laura Rosane; da HORA, Dermeval. As consoantes do português. *In:* BISOL, Leda (org.). *Introdução a estudos de fonologia do português brasileiro – 4a edição.* Porto Alegre: EDIPUCRS, 2005, p. 207-241.

OLIVEIRA, Dermeval da Hora. *A palatalização das oclusivas dentais: variação e representação não-linear*. 292 ff. Tese: Doutorado em Letras. PUCRS. Porto Alegre, 1990.

PRESTES, Susana Pinheiro da Cruz. Análise Acústica da Produção de Consoantes Oclusivas Iniciais por Falantes Nativos de PB. *Anais do X encontro do Círculo de Estudos Linguísticos do Sul.* Cascavel: UNIOESTE, 2012.

REIS, Mara; NOBRE-OLIVEIRA, Denize**.** Effects of perceptual training on the identification and production of English voiceless plosives aspiration by Brazilian EFL learners. *Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech.* Florianopolis: UFSC, 2008. p. 372-381.

SILVA, Thaïs Cristófaro. *Pronúncia do Inglês para Falantes do Português Brasileiro.* Contexto, 1ª edição, 2012.

YAVAS, Mehmet**.** Factors influencing the VOT of English long lag stops and interlanguage phonology. In: RAUBER, Andrea S.; WATKINS, Michael A.; BAPTISTA, Barbara O. (eds.). New Sounds 2007: *Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech.* Florianópolis: UFSC, 2008, p. 492-498.

_____; WILDERMUTH, Renée. *The effects of place of articulation and vowel height in the acquisition of English aspirated stops by Spanish speakers. IRAL*, v. 44, p. 251-263, 2006.

ZIMMER, Márcia Cristina. *A transferência do conhecimento fonético-fonológico do português brasileiro (L1) para o inglês (L2) na recodificação leitora: uma abordagem conexionista.* Tese (Doutorado em Letras) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, 2004.

_____; ALVES, Ubiratã Kickhöfel. *On the Status of Terminal Devoicing as an Interlanguage Process among Brazilian learners of English*. Ilha do Desterro, v. 55, p. 41-62, 2008.

_____; SILVEIRA, Rosane; ALVES, Ubiratã Kickhöfel. *Pronunciation instruction for Brazilians: bringing theory and practice together*. Newcastle upon Tyne: Cambridge Scholars Publishing, 2009.

_____; ALVES, Ubiratã Kickhöfel. *Learning to orchestrate time: Voicing patterns and gestural drift in L2 speech production.* Livro de Resumos do São Paulo School of Advanced Studies in Speech Dynamics. São Paulo, 2010, p. 47-48.

_____; ALVES, Ubiratã Kickhöfel. Uma visão dinâmica da produção da fala em L2: o caso da dessonorização terminal. *Revista da Abralin*, número especial 2, 2012.

# APPENDICES

## APPENDIX 1 – FORMULÁRIO DE CONSENTIMENTO

**Prezado Informante:**
Por favor, leia o texto a seguir. Ele apresenta informações importantes a respeito do estudo de que você fará parte. Após isso, assine o documento, indicando que você entende a natureza desta pesquisa e que você consente participar dela.

**DADOS DA PESQUISA**

**Investigadores:**
Prof. Dr. Ubiratã Kickhöfel Alves (UFRGS)
Bruno M. Schwartzhaupt (UFRGS)

**PROPÓSITO E BENEFÍCIOS**

Este estudo tem por objetivo investigar a produção de um aspecto fonético em palavras do português e do inglês, produzidas por falantes nativos do português brasileiro. A pesquisa visa a contribuir com estudos referentes à aquisição de língua inglesa como língua estrangeira.

**PROCEDIMENTOS**

O informante realizará uma tarefa de produção de palavras do Inglês e do Português Brasileiro. Nessa tarefa, o participante realizará a leitura oral de palavras em língua nas duas línguas. As palavras serão apresentadas individualmente, em um computador do tipo *laptop*, cabendo ao participante pronunciá-las. Previamente a essa tarefa, o informante será submetido a uma verificação de nivelamento em L2 (Oxford Placement Test 1), com vistas a identificar o seu nível de proficiência para posterior enquadramento na pesquisa. O áudio da tarefa será gravado e armazenado em arquivos do tipo .wav, no computador utilizado nas sessões de coleta.

**OUTRAS INFORMAÇÕES**

A participação no estudo é de caráter voluntário. Todos os participantes têm a liberdade de cancelar a participação a qualquer momento. O material de áudio coletado será ouvido apenas pelo professor envolvido na pesquisa e orientando. Somente eles terão acesso ao material de áudio. As gravações em áudio ficarão de posse dos pesquisadores, e os dados coletados poderão ser utilizados em estudos posteriores. A identidade de todos os participantes permanecerá confidencial - ao participar da pesquisa, o informante receberá um Número de Identificação, de modo que o nome do participante não seja nunca divulgado.

**DECLARAÇÃO**

Declaro que li e compreendi as informações acima e que consinto participar desta pesquisa.

..... ..... .....
Nome Assinatura Data

APPENDIX 2 – FICHA DE INFORMAÇÕES DO PARTICIPANTE

Por favor, preencha o formulário abaixo. Sua participação neste estudo é muito importante. Obrigado!

1)Nome: .....

2) Data de nascimento e idade: .....

3) Cidade natal: .....

4) Cidade natal do pai: .....

5) Cidade natal da mãe: .....

6) Instituição de ensino em que foi contatado.....

7) Nível de adiantamento ( ____ semestre / básico / intermediário, etc.):

8) Carga horária semanal de língua inglesa dentro da instituição: .....

9) Carga horária semanal de língua inglesa em caráter extracurricular: .....

10) Escolaridade:
( ) ensino médio ( ) superior incompleto ( ) superior completo ( ) pós-graduação

11) Línguas adquiridas até os 6 anos de idade: .....

12) Idade com que iniciou a estudar inglês: .....

13) Cidade de residência e período de tempo em que reside nesta cidade: .....

14) Instituição em que iniciou a estudar inglês: .....

15) Outras línguas que você pode falar, além do português e do inglês: .....

16) Países de língua inglesa que já visitou: .....

17) Período de tempo da visita: .....

18) Países de língua inglesa em que já residiu: .....

19) Período de tempo em que residiu: .....

20) Somando-se todos os períodos em que você estudou a língua inglesa em cursinho de línguas / intercâmbio / universidade, qual é o tempo total: ..... anos e ..... meses.

Data: .....
Informante nº: .....

APPENDIX 3 – CONSENT FORM

**Dear Participant:**
Please read the following text. It provides important information regarding the study of which you are about to take part. After doing so, please sign this document, indicating that you understand the purposes of this study and that you consent to participate in it.

**Investigators:**
Prof. Dr. Ubiratã Kickhöfel Alves (UFRGS)
Bruno Moraes Schwartzhaupt (UFRGS)

**PURPOSE OF STUDY**

This study investigates the production of a phonetic feature by native speakers of American English and Brazilian learners of English. The research aims at contributing with studies related to the acquisition of English as a Foreign Language.

**PROCEDURES**

The participant will perform a task in which they read aloud several English words. The target-words will be presented individually in a *laptop*. The áudio recorded in this task will be saved in .wav files in the computer used in it.

**OTHER INFORMATION**

Participation in this study is voluntary. All participants are free to cancel their participation whenever they wish. The audio here recorded will only be heard by the professor involved in the research and his assistant. They will be the only ones to have access to these files. The recordings will remain under the possession of the researchers in question, and they can use such files in posterior studies. The participants' identity will remain confidential: by participating in this study, the informant will be given na Identification Number, so their name will not be revealed.

# DECLARATION

I declare that I have read and understood the information presented above, and that I consent to participate in this study.

..... ..... .....
Nome Assinatura Data

APPENDIX 4 – PARTICIPANT INFORMATION FORM

Please fill in the form below. Your participation in this study is very important. Thank you.

1) Name: .....

2) Date of birth and age: ..... …….

3) Hometown: .....

4) Father's hometown: .....

5) Mother's hometown: .....

6) Institution in which you were contacted .....

7) Education:
( ) high school ( ) undergraduate ( ) graduate ( ) post-graduate

8) Languages acquired before reaching 6 years of age: .....

9) Age in which you started learning another language: .....

10) Length of time studying a foreign language: ……

11) Hours you dedicate to the learning of a second language weekly (which language?) .....

12) City of residence / length of time living in this city: .....

13) Foreign countries you have visited: .....

14) Length of time you stayed in this country: .....

15) Foreign countries you have lived in: .....

16) Length of time you lived in this country: .....

Date: ….....
Participant nº: ….....