



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS

Gustavo Duarte Régis

**Implicit and explicit phonological awareness of English VOT in
Brazilian learners of English**

Florianópolis

2026

Gustavo Duarte Régis

**Implicit and explicit phonological awareness of English VOT in
Brazilian learners of English**

Dissertação submetida ao Programa de Pós-Graduação
em Inglês da Universidade Federal de Santa Catarina
como requisito parcial para a obtenção do título de
Mestre em Estudos Linguísticos.

Orientadora: Prof^a. Hanna Kivistö-de Souza, Dr^a.

Florianópolis
2026

Régis, Gustavo

Implicit and explicit phonological awareness of English
VOT in Brazilian learners of English / Gustavo Régis ;
orientadora, Hanna Kivistö-de Souza, 2026.
123 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro de Comunicação e Expressão, Programa de
Pós-Graduação em Inglês: Estudos Linguísticos e
Literários, Florianópolis, 2026.

Inclui referências.

1. Inglês. 2. Aquisição da fala em L2. 3. Hipótese da
interface. 4. Conhecimento metalinguístico. 5. Consciência
fonológica. I. Kivistö-de Souza, Hanna. II. Universidade
Federal de Santa Catarina. Programa de Pós-Graduação em
Inglês: Estudos Linguísticos e Literários. III. Título.

Gustavo Duarte Régis

**Implicit and explicit phonological awareness of English VOT in
Brazilian learners of English**

O presente trabalho em nível de Mestrado foi avaliado e aprovado, em 25 de fevereiro de 2026, pela banca examinadora composta pelos seguintes membros:

Prof^a. Hanna Kivistö-de Souza, Dr^a.
Universidade Federal de Santa Catarina

Prof. Germán Zárate-Sández, Dr.
Western Michigan University

Prof. Felipe Flores Kupske, Dr.
Universidade Federal do Rio Grande do Sul

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestre em Estudos Linguísticos.



Documento assinado digitalmente
Alinne Balduino Pires Fernandes
Data: 02/03/2026 15:13:10-0300
CPF: ***.588.801-**
Verifique as assinaturas em <https://v.ufsc.br>

Coordenação do Programa de Pós-Graduação



Documento assinado digitalmente
Hanna Emilia Kivisto de Souza
Data: 02/03/2026 13:30:54-0300
CPF: ***.442.279-**
Verifique as assinaturas em <https://v.ufsc.br>

Prof^a. Hanna Kivistö-de Souza, Dr^a.
Orientadora

Florianópolis, 2026.

In memory of Catarina Kasten.

ACKNOWLEDGEMENTS

I would like to take a moment to express my gratitude to the dearest and most important people.

First, I would like to thank Professor Hanna, who very cheerfully led this work with me. Thank you for sharing your expertise and for dedicating your time and effort to improving this study. Researching awareness became fascinating under your supervision.

Gostaria de agradecer à minha família, pais, irmã, avós, bisa, tias e tios, primas e primos. Obrigado, mãe e pai, por me apoiarem, não apenas nesse trabalho, mas na vida. Vocês são minha inspiração e motivação, eu amo tanto vocês. Muito obrigado, Dada. Obrigado por ser irmã, no sentido mais bonito da palavra. Você significa muito para mim. Vocês três não precisariam fazer nada e escolheram fazer tudo. Vocês me fizeram entender o que é amar uma criança, não só uma, mas duas. Laura e João, o dindo ama vocês infinito e para sempre! Eu tenho um grande motivo para ser grato todos os dias que é ter vocês comigo.

Além da minha família, eu sou grato às grandes amigas e amigos que completam o sentido de tudo.

Obrigado, Bia, Maria e Isa, por estarem comigo desde muito antes, vocês são muito mais para mim do que podem imaginar. Uma parte substancial das minhas melhores e maiores lembranças de momentos triviais e importantes é com vocês. Desejo estar sempre juntos. Vocês são mais valiosas que qualquer diamante que procurávamos na cachoeira. Eu amo vocês.

Além da infância, fui presenteado com amizades fundamentais na escola e na faculdade. Obrigado Ceci, Fran, Je e Alanna. Fran, desde o sétimo ano eu tenho a sorte de poder te irritar. Ceci, sempre lembro das nossas conversas e músicas da sua *playlist* voltando dos rolês pra te deixar em casa. Je, eu falo de ti para todos - seu jeito me fascina. Alanna, eu amo muito fazer trilhas, mas amo muito mais quando é com você. Obrigado por me motivarem tanto de uma forma tão espontânea.

I couldn't be more grateful for my closest friends from Letras, who are pivotal to the development of this work. THANK YOU, Janete, Luiza, Mariana, and Maysa. You are a big part of the reason I fell in love with Linguistics, an area I had never thought I would engage with. You are the most special literary girls who made me understand the joy of reading the books you like and the authors you love (and the movies you watch). I admire you so much. I will drive Janete's hill as many times as needed to laugh with you and learn from you. Because you are the intelligent people who impressed (and continue impressing) me at UFSC.

Not only in my undergraduate studies did I meet incredible people, but fortunately, during my master's as well. I want to thank my two closest NUPFFALE friends, Manu and Crys. Thank you for sharing this experience with me. I am excited for our next step.

During this journey, I found a special 'mate'. Sam, eu te amo. Você é o meu carioca favorito. Todas as suas ligações durante meu horário de estudo não atrapalharam, ajudaram.

Apart from friends, I am also grateful for the people and institutions that are part of this work. I want to thank Professor Rosane for facilitating our participation in the academic world with her patience and kindness. I am also grateful to Professors Felipe Kupske and Germán Zárate-Sández for participating in my exam board and for their willingness to help. I also thank Universidade Federal de Santa Catarina (UFSC) and the Programa de Pós-Graduação em Inglês (PPGI), as well as CAPES, for making this study possible through the grant.

Finally, I am grateful to the friends who directly contributed to my study by sharing it online, with family and friends, and to the participants who voluntarily spent their time completing the data collection session. Thank you!

ABSTRACT

Despite the heated theoretical discussion on the possible interface between implicit and explicit second language (L2) knowledge, few studies have empirically analyzed the relationship between these two types of knowledge or awareness (Roehr-Brackin, 2024), and the ones that have done so, analyzed morphosyntactic but not phonological knowledge (Kim; Godfroid, 2023; Suzuki; DeKeyser, 2017). Implicit and explicit L2 phonological knowledge constitute what is known as ‘L2 phonological awareness’, in which explicit awareness refers to conscious and potentially verbalizable knowledge, and implicit awareness refers to subconscious and unverbalizable knowledge. The general aim of the present study is to empirically investigate implicit and explicit L2 phonological awareness. For that, knowledge about English Voice Onset Time (VOT) in Brazilian learners of English was analyzed with three objectives in mind: (1) to explore the extent to which Brazilian learners of English possess implicit and explicit awareness about English VOT; (2) to examine the relation between implicit and explicit L2 phonological awareness; (3) to examine the variance in implicit L2 phonological awareness about VOT in terms of L2 proficiency, L2 input, and explicit L2 phonological awareness. Brazilian learners of English ($N = 37$) completed four instruments virtually: a background questionnaire, a perception test based on Flege and MacKay (2004) containing items with accurate and inaccurate VOT intended to measure implicit awareness, a novel questionnaire developed to assess explicit awareness through audio and text, and a vocabulary measure to estimate participants’ English proficiency. Results revealed that Brazilian learners manifested generally low levels of implicit and explicit awareness of English VOT. The correlation between implicit and explicit L2 phonological awareness did not reach statistical significance, although the effect size was close to medium (.26). Moreover, explicit L2 phonological awareness and L2 proficiency, but not L2 input, were found to predict implicit L2 phonological awareness significantly. The findings contribute to our understanding of the nature of implicit and explicit L2 knowledge and have pedagogical implications for the teaching of English to Brazilians. Identifying the gaps in Brazilian learners’ knowledge of English enables educators to focus on aspects of the language that may be lacking attention. Ultimately, the development of an instrument designed to measure explicit L2 phonological awareness can inform future research in this area.

Keywords: L2 speech acquisition; interface hypothesis; metalinguistic knowledge.

RESUMO

Apesar do intenso debate teórico sobre a possível interface entre o conhecimento implícito e explícito da segunda língua (L2), poucos estudos analisaram empiricamente a relação entre esses dois tipos de conhecimento ou consciência (Roehr-Brackin, 2024), e os que o fizeram analisaram o conhecimento morfossintático, mas não fonológico (Kim; Godfroid, 2023; Suzuki; DeKeyser, 2017). Conhecimento fonológico implícito e explícito em L2 constitui o que é conhecido como ‘Consciência fonológica em L2’, em que consciência explícita refere-se ao conhecimento consciente e potencialmente verbalizável e consciência implícita ao conhecimento subconsciente e não verbalizável. O objetivo geral deste estudo é investigar empiricamente o conhecimento fonológico implícito e explícito em L2. Para isso, o conhecimento sobre o *Voice Onset Time* (VOT) do inglês em aprendizes brasileiros de inglês foi analisado com três objetivos em mente: (1) explorar em que medida os aprendizes brasileiros de inglês possuem consciência implícita e explícita sobre o VOT do inglês; (2) examinar a relação entre a consciência fonológica implícita e explícita da L2; (3) examinar a variância na consciência fonológica implícita da L2 sobre o VOT em termos de proficiência na L2, input da L2 e consciência fonológica explícita da L2. Trinta e sete aprendizes brasileiros de inglês responderam virtualmente a quatro instrumentos: um questionário de informações gerais, um teste de percepção baseado em Flege e MacKay (2004) contendo itens com VOT correto e incorreto com o objetivo de mensurar a consciência implícita, um novo questionário desenvolvido para avaliar a consciência explícita por meio de áudio e texto, e um teste de vocabulário para estimar a proficiência em inglês dos participantes. Os resultados revelaram que os aprendizes brasileiros manifestaram baixos níveis de consciência implícita e explícita do VOT do inglês. A correlação entre a consciência fonológica implícita e explícita da L2 não atingiu significância, embora o tamanho do efeito tenha sido próximo ao tamanho médio (.26). Além disso, a consciência fonológica explícita da L2 e a proficiência na L2, mas não o input da L2, foram preditoras significativas da consciência fonológica implícita da L2. Os achados contribuem para a compreensão da natureza do conhecimento implícito e explícito da L2 e têm implicações pedagógicas para o ensino de inglês para brasileiros. Identificar as lacunas no conhecimento de inglês dos aprendizes brasileiros permite que educadores se concentrem em aspectos da língua que podem estar recebendo menos atenção. Por fim, o desenvolvimento de um instrumento destinado a medir a consciência fonológica explícita da L2 pode orientar pesquisas futuras na área.

Palavras-chave: aquisição da fala em L2; hipótese da interface; conhecimento metalinguístico.

LIST OF FIGURES

Figure 1 – The spectrogram and waveform of the words <i>tip</i> and <i>dip</i>	20
Figure 2 – Voice Onset Time categories	21
Figure 3 – Commonest production of stop consonants in Portuguese and English	27
Figure 4 – The Manipulation Editor window inside Praat showing the duration function (bottom frame) for the target word <i>keep</i> by a male speaker	40
Figure 5 – A test trial of the <i>Error Detection Test</i>	42
Figure 6 – The two instruction windows of the <i>Error Detection Test</i>	42
Figure 7 – Text presented before the questionnaire measuring explicit awareness	43
Figure 8 – Screenshot of the <i>Text-based section</i> with the first set of statements	46
Figure 9 – Screenshots of the <i>Text-based section</i> with the second set of statements	48
Figure 10 – A screenshot showing one item of the <i>Audio-based section</i>	50
Figure 11 – Boxplot showing the percentage of correct responses per voiceless stop from the <i>Error Detection Test</i>	61
Figure 12 – Boxplot showing the percentage of correct responses to accurate and inaccurate tokens from the <i>Error Detection Test</i>	62
Figure 13 – Boxplot showing the <i>Implicit Awareness Score</i> and the <i>Explicit Awareness Score</i>	62
Figure 14 – Boxplot for the percentage of correct responses per item group (BP accurate and inaccurate, and English accurate and inaccurate) from the <i>Audio-based section</i>	64
Figure 15 – Boxplot for the percentage of correct responses to the first and second sets of statements from the <i>Text-based section</i>	64
Figure 16 – Frequency of the certainty ratings for each item type according to participants' correct and incorrect responses.....	66
Figure 17 – Scatter plot of the correlation between the <i>Implicit Awareness Score</i> and the <i>Explicit Awareness Score</i>	74

LIST OF TABLES

Table 1 – Differences between implicit and explicit knowledge.....	7
Table 2 – Mean VOT values in ms for the voiceless and voiced stops in English.....	22
Table 3 – Mean VOT values in ms for the voiceless and voiced stops in PB	25
Table 4 – Participants’ background information	34
Table 5 – Number of participants and their proportion of Brazilian Portuguese (BP) and English use across their 5 closest persons	36
Table 6 – Target words and distractors included in the <i>Error Detection Test</i>	38
Table 7 – Target words and distractors included in the <i>Text-based section</i>	47
Table 8 – Words included in the <i>Audio-based section</i>	50
Table 9 – Descriptive statistics of <i>Implicit Awareness Score</i> and <i>Explicit Awareness Score</i> .	60
Table 10 – Percentage of correct responses to the accurate and inaccurate items from the <i>Audio-based section</i>	63
Table 11 – Frequency and relative frequency (in percentage) of the certainty ratings for each item type according to participants’ correct and incorrect responses and for all responses together	65
Table 12 – Open-ended responses to the first and second sets of statements related to stops from the <i>Text-based section</i> of the questionnaire measuring explicit awareness.....	67
Table 13 – Correlation coefficients between the variables ($N = 36$)	76
Table 14 – Results from the multiple regression analysis ($N = 36$).....	77

TABLE OF CONTENTS

1 INTRODUCTION	1
2 REVIEW OF LITERATURE.....	4
2.1 LANGUAGE AWARENESS AND LANGUAGE INPUT	4
2.1.1 Implicit and explicit knowledge	5
2.1.2 The relation between implicit and explicit knowledge.....	7
2.1.3 L2 phonological awareness	11
2.1.4 L2 input quantity and quality	15
2.2 VOICE ONSET TIME (VOT).....	18
2.2.1 English VOT	21
2.2.2 Brazilian Portuguese VOT	24
2.2.3 Acquisition of English VOT by Brazilians.....	27
3 THE PRESENT STUDY	32
3.1 OBJECTIVES AND RESEARCH QUESTIONS	32
3.2 PARTICIPANTS	33
3.3 INSTRUMENTS.....	36
3.3.1 Implicit L2 phonological awareness	37
3.3.1.1 Stimuli	37
3.3.1.2 Stimuli recording	38
3.3.1.3 Stimuli treatment	39
3.3.1.4 Test design.....	41
3.3.2 Explicit L2 phonological awareness.....	43
3.3.2.1 Text-based section	45
3.3.2.2 Audio-based section	49
3.3.3 Proficiency test.....	51
3.3.4 Background Questionnaire.....	52
3.4 PROCEDURES.....	53
3.5 DATA ANALYSIS.....	54
3.5.1 Implicit awareness score	54
3.5.2 Explicit awareness score	55
3.5.3 L2 input score	58
3.5.4 L2 proficiency score	59

3.5.5 Research questions	59
4 RESULTS AND DISCUSSION	60
4.1 RQ1: BRAZILIAN LEARNERS' LEVELS OF IMPLICIT AND EXPLICIT AWARENESS OF ENGLISH VOT.....	60
4.1.1 Discussion of RQ1.....	70
4.2 RQ2: THE RELATION BETWEEN IMPLICIT AND EXPLICIT KNOWLEDGE.....	72
4.2.1 Discussion of RQ2.....	74
4.3 RQ3: VARIANCE IN IMPLICIT L2 AWARENESS ABOUT VOT	75
4.3.1 Discussion of RQ3.....	78
4.4 GENERAL DISCUSSION	80
5 CONCLUSIONS.....	82
5.1 LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH.....	84
5.2 PEDAGOGICAL IMPLICATIONS	85
 REFERENCES	 88
 APPENDIX A – TOKENS INCLUDED IN THE FAMILIARIZATION SESSION OF THE <i>ERROR DETECTION TEST</i>	 97
APPENDIX B – TEST TOKENS INCLUDED IN THE <i>ERROR DETECTION TEST</i> ...	98
APPENDIX C – BACKGROUND QUESTIONNAIRE	99
APPENDIX D – INTERNAL CONSISTENCY OF THE QUESTIONNAIRE MEASURING EXPLICIT AWARENESS.....	106
APPENDIX E – FREQUENCY AND RELATIVE FREQUENCY OF THE CERTAINTY-RATINGS.....	107
APPENDIX F – DATA ANALYSIS OF THE <i>L2 INPUT SCORE</i>.....	108
APPENDIX G – INDIVIDUAL SCORES.....	109

1 INTRODUCTION

Learning a second language (L2)¹ has become an ordinary experience in today's globalized world. Language learners develop different competencies in the target language to guarantee they can communicate well with others. Speaking, particularly, demands the development of different language skills, and a foreign accent is common in L2 speakers' speech. Nevertheless, the occasional mispronunciation of some L2 features can lead to communication breakdowns and sometimes even embarrassment (Lima Jr., 2015), for example, the replacement of a consonant by another, resulting in a different word, such as 'tie' /tai/ and 'die' /dai/ when pronounced as homophones. Therefore, researchers have sought to investigate the features posing difficulty in L2 speech and the factors influencing L2 speech acquisition.

Features analyzed in L2 speech can be roughly categorized as segmental (vowels and consonants) and suprasegmental (e.g., intonation and rhythm). In segmental studies, the differences in the phoneme inventory between languages and the problems arising from it are addressed. One such difference, for example, is the variation in Voice Onset Time (VOT), which refers to voicing and release duration in the production of stop consonants (Yavas, 2011). When languages possess differing VOT patterns, such as the case of Brazilian Portuguese (BP) and American English², opportunities for fruitful studies in the field of second language acquisition (SLA) arise.

In addition to that, factors like language use and language awareness have been investigated in relation to L2 speech acquisition. A large number of studies have found language use to play a significant role in L2 speech acquisition (e.g., Flege; Munro; MacKay, 1995; Flege; Frieda; Nozawa, 1997; Moyer, 2009). To that end, it appears that L2 learners who use the L2 more often are more likely to display higher levels of L2 speech accuracy than those L2 learners who seldom use the language (Piske; MacKay; Flege, 2001). Moreover, a still-growing body of research has been interested in the role played by language awareness, that is, "explicit knowledge about language, and conscious perception and sensitivity in language learning, language teaching and language use" (Association for Language Awareness, 2025). Therefore, language awareness is constituted of two types of knowledge: encyclopedic knowledge available for conscious reflection, also known as *explicit knowledge*, and intuitive, unconscious

¹ In the present study, the terms 'second language', 'additional language', and 'foreign language' will be regarded as the same construct, which, in general terms, refers to a language learned after early childhood.

² Hereafter, when 'English' is mentioned without reference to a specific variety, it refers to the General American English.

knowledge, also referred to as *implicit knowledge* (Kivistö-de Souza, 2015). Findings suggest that higher levels of L2 awareness contribute to L2 speech acquisition (Baker; Trofimovich, 2006; Kennedy; Trofimovich, 2010; Teló; Kivistö-de Souza, 2023).

The study of language awareness is highly relevant as it helps understand two major issues in SLA: how languages can be better learned and taught, and how L2 acquisition differs from first language (L1) acquisition (Kim, 2020). Krashen (1981, 1985) proposed that L2 acquisition is much the same as L1 acquisition. In this view, the L2 can be simply picked up from exposure, and formal instruction does not contribute much to language fluency. For Krashen (1981, 1985), formal instruction facilitates only the acquisition of explicit language knowledge, the one that is often inaccurate and effortful to access, while implicit language knowledge is acquired through natural, meaningful communication, that is, subconsciously, and one type of knowledge does not relate to the other. On the contrary, Schmidt (1995, 2010) argued that just being exposed to the target language is not enough for the L2 learner to achieve high levels of proficiency, as consciousness is believed to play a major role in L2 acquisition. Building on this position, Ellis (2006) argued that implicit knowledge is related to explicit knowledge, and so, language learners benefit from formal instruction as it may help bring consciousness to specific features of language that may go unnoticed. However, the relation between the two types of language knowledge is not yet established. Despite the heated theoretical debate, few studies have sought to empirically investigate the relationship between implicit and explicit language knowledge (Kim; Godfroid, 2023; Suzuki; DeKeyser, 2017).

Therefore, the primary objective of the present study is to examine the relationship between implicit and explicit knowledge about L2 phones, in other words, between implicit and explicit L2 phonological awareness. At first, the present study aims to explore the extent to which Brazilian learners of English possess implicit and explicit knowledge about English VOT, given that English and BP possess different VOT patterns. Second, it aims to analyze the relationship between implicit and explicit L2 phonological awareness of VOT. Finally, the effect of L2 use, L2 proficiency, and explicit L2 phonological awareness on the variation of the implicit phonological awareness is investigated.

Considering the relevance of L2 pronunciation in communication, a better understanding of the complexity of phonological awareness within SLA is highly relevant. This study intends to contribute to an emerging area within language awareness studies by obtaining empirical insight into the relationship between implicit and explicit L2 phonological awareness. The findings may help us understand if the association found between the two types of morphosyntactic L2 knowledge (Kim; Godfroid, 2023; Suzuki; DeKeyser, 2017) can also be

seen in phonology, as no studies to the best of our knowledge have investigated that. Moreover, investigating phonological awareness about English VOT by Brazilian learners of English may help English teachers be aware of topics worth teaching their Brazilian students by realizing what is likely to contribute to their previous knowledge in phonology, given that phonological awareness is believed to affect pronunciation (Kennedy; Trofimovich, 2010; Kivistö-de Souza, 2015; Ramírez Verdugo, 2006). What is more, the present study ventured to create an instrument intended to measure explicit L2 phonological awareness of VOT, as, to the best of my knowledge, no previous studies have presented an instrument to do so. By creating this instrument, we may contribute to the significant deficiency in SLA research that is instruments measuring L2 knowledge (Roehr-Brackin, 2024), especially knowledge about L2 phonology (Ellis, 2005). Furthermore, L2 phonological awareness remains an under-researched area within SLA, and more evidence is necessary to draw more concrete conclusions on the topic (Kivistö-de Souza, 2017).

The present thesis is divided into five chapters. This first chapter, ‘Introduction’, has presented the main constructs and objectives of the study. The second and following chapter, ‘Review of literature’, will provide the theoretical foundation of the study in two main sections, ‘Language awareness and language input’ and ‘Voice Onset Time (VOT)’. The third chapter, ‘The present study’, is divided into five sections. The first section announces the objectives and research questions of the study. The second and third sections display the participants and instruments, respectively. The fourth section, ‘Procedures’, presents the steps to conduct the study, while the fifth section, ‘Data analysis’, presents the analysis of each instrument separately, along with the statistical procedures to answer the research questions. The fourth chapter, ‘Results and discussion’, reports the results for each research question and presents a discussion of the findings. Finally, the fifth chapter regards the ‘Conclusions’, in which concluding remarks to the thesis are presented along with limitations and suggestions for future research and pedagogical implications.

2 REVIEW OF LITERATURE

This chapter will provide the background for the study in two major parts. The first part addresses the broad topics of language awareness and language input, where the differences and relationship between implicit and explicit knowledge will be explored. A discussion on L2 phonological awareness and the role of L2 input quantity and quality on linguistic dimensions follows this. The second part refers to Voice Onset Time (VOT), the phonetic feature analyzed in the present study. This part will address the English and Brazilian Portuguese VOT patterns as well as the acquisition of English VOT by Brazilians.

2.1 LANGUAGE AWARENESS AND LANGUAGE INPUT

Learners of a new language tend to have difficulties not only in pronouncing but also in *hearing* the new sounds (Jenkins; Strange; Polka, 1995). The differences between the L1 and L2 sounds and the ways they are employed to convey meaning in each language seem to affect how language learners perceive and produce L2 sounds (Zimmer; Silveira; Alves, 2009). In addition to that, individual differences, such as Age of Onset of Learning (AOL) and motivation to learn an L2, have been demonstrated to affect L2 speech acquisition (Piske; MacKay; Flege, 2001). In this regard, it appears that younger AOL and higher levels of motivation benefit L2 speech development (e.g., Lima Jr., 2012; Moyer, 1999). Besides age-related and sociopsychological factors, studies have also been interested in cognitive factors affecting L2 speech acquisition. Within such studies, phonological awareness, that is, learners' general knowledge about the target language phonology, is analyzed and shown to likely matter to L2 speech development (e.g., Baker; Trofimovich, 2006; Kennedy; Trofimovich, 2010).

Phonological awareness is a field of study within the broader area of language awareness. Language awareness refers to learners' knowledge about the target language in general, not only in phonology, and is more specifically defined by the Association for Language Awareness (2025) as “explicit knowledge about language, and conscious perception and sensitivity in language learning, language teaching and language use”. Following that, language awareness consists of two types of knowledge: explicit knowledge, which is conscious and potentially verbalizable knowledge, and implicit knowledge (also framed as *sensitivity* to language), which is subconscious and unverbalizable knowledge. However, many studies focus on the explicit part of language awareness (Kivistö-de Souza, 2015) and refer to this construct as *metalinguistic knowledge* (or metalinguistic awareness), which relates to “learner's explicit

knowledge about the syntactic, morphological, lexical, phonological and pragmatic features of the L2” (Roehr; Gánem-Gutiérrez, 2009, p. 165). Thereupon, language awareness comprises both metalinguistic knowledge and implicit knowledge³ about language.

2.1.1 Implicit and explicit knowledge

Implicit and explicit knowledge are two different constituents of language awareness. These two types of knowledge are argued to be stored in different parts of the brain and considered separate but connected systems that are not completely distinct in performance (Ellis, 2008).

Implicit or proceduralized knowledge refers to the unconscious knowledge used in automatized actions. This tacit knowledge is attributed to intuition and therefore cannot be verbalized by the speaker since any attempt to verbalize implicit knowledge involves the creation of an explicit representation first (Ellis, 2005). In Ellis (2004, p. 231), implicit knowledge is referred to as the “basic linguistic competence”. Common examples drawing on implicit knowledge are instantly recognizing accented speech and ungrammatical sentences (Ellis, 2005).

Explicit or declarative knowledge is the conscious knowledge that speakers know they possess. This type of knowledge does not differ from encyclopedic knowledge of any other kind (Ellis, 2004), e.g., knowing what the capital city of Brazil is, or which numbers are prime. Given its conscious nature, explicit knowledge can be verbalizable, although it does not depend on verbalization to exist (Ellis, 2004). In this regard, speakers can use or not use metalanguage when verbalizing explicit knowledge about language, e.g., describing a grammatical rule (Ellis, 2004). Ellis (2004, p. 229) defined explicit knowledge about language as “the conscious awareness of what a language or language in general consists of and/or of the roles that it plays in human life”.

Implicit and explicit knowledge are believed to be generated from different educational experiences (Ellis, 2008). While implicit knowledge is acquired by everyday language use, explicit knowledge is established by consciously registering the input (Ellis, 2005). Accordingly, implicit knowledge is the result of implicit learning, that is, engaging in natural, simple, and unconscious activities such as natural, meaningful communication, in contexts where target language input is abundant and complex (Ellis, 2008). Explicit knowledge, in turn,

³ In this study, ‘implicit knowledge’ and ‘implicit awareness’, and ‘explicit knowledge’ and ‘explicit awareness’ will be used interchangeably to refer to the same construct.

results from explicit learning, “a more conscious operation where the individual attends to particular aspects of the stimulus array and volunteers and tests hypotheses in a search for structure” (Ellis, 2008, p. 3).

On that account, explicit knowledge develops ontogenetically later than implicit knowledge and does not suffer from age constraints like implicit knowledge (Ellis, 2008). Evidence that implicit knowledge develops before explicit knowledge in life is seen in the acquisition of L1 grammar. By engaging in everyday communication, children automatically acquire knowledge of their L1 structure, but they will be able to verbalize their knowledge only after going through some reflection, formal education, or another event that brings language into consciousness (Ellis, 2008). Besides that, children seem better at picking up language patterns through communication, whereas adult language learners seem to benefit more from explicit language instruction. This supports the belief that implicit knowledge is affected by age constraints, while explicit knowledge is not (Ellis, 2005). Therefore, languages are likely better implicitly learned at younger ages, even though implicit learning mechanisms remain available throughout life.

Explicit and implicit knowledge are stored in different parts of the brain and involve different types of processes (Ellis, 2004). Briefly speaking, implicit knowledge is located in various areas of perceptual and motor cortex, and explicit knowledge in the hippocampus and related limbic structures (Ellis, 2009). One important characteristic that differentiates the two types of knowledge is how they are retrieved or processed. Implicit knowledge is available for rapid, automatic use and is considered more accurate, whereas explicit knowledge takes longer to be retrieved due to its controlled character and is frequently inaccurate or imprecise (Ellis, 2004, 2005). Yet, some researchers (e.g., DeKeyser, 2003) argue that implicit knowledge may become proceduralized to a point where it becomes hard to distinguish between proceduralized explicit knowledge and implicit knowledge (further discussion in *Section 2.1.2*).

As a result, different tasks are employed to assess each type of knowledge (Kivistö-de Souza, 2015). Due to the automatic retrieval of implicit knowledge, tasks aiming at tapping into this type of knowledge are commonly timed in a way that task takers do not have the time to retrieve conscious knowledge but depend on intuition. On the other hand, tasks examining explicit knowledge usually address the subject more directly, which means that learners can access their conscious knowledge. Moreover, language tasks considered difficult by the learner may naturally lead to an effort to exploit explicit knowledge (Ellis, 2004). However, as Ellis (2005) noted, it is impossible to create a task drawing on one specific type of knowledge only, and there is no guarantee that task takers will employ the intended type of knowledge, since

they will use whatever source is available to accomplish the task. This is because the two types of knowledge are believed to interact at the level of performance despite being well distinguished at the level of learning (Ellis, 2004).

Table 1 provides a summary of the differences between implicit and explicit knowledge across the different characteristics.

Table 1 – Differences between implicit and explicit knowledge

Characteristic	Implicit knowledge	Explicit knowledge
Cerebral location	neocortex	hippocampus
Access	automatic and fast	controlled and slower
Type of knowledge	intuitive	encyclopedic
Verbalizable?	no	yes, to different degrees
Role of awareness	unconscious	conscious
Order of acquisition	1	2
Age-constrained?	yes	no

Source: adapted from Kivistö-de Souza (2015, p. 19).

2.1.2 The relation between implicit and explicit knowledge

Although there is a general agreement on the definitions of implicit and explicit knowledge, a strong theoretical divergence permeates the field of SLA regarding the relation between these two types of knowledge. Questions arise about whether explicit knowledge can become implicit or the other way round. Or, yet, whether explicit knowledge can at least facilitate the acquisition of implicit knowledge, and vice versa (Suzuki; DeKeyser, 2017). These questions are of great relevance as it has been argued that implicit and explicit knowledge may derive from different educational experiences (Ellis, 2008). In this way, explicit learning (e.g., formal instruction) facilitates mainly the acquisition of explicit knowledge, and thus, the majority of L2 learners who attend formal language instruction are deemed to have developed mainly explicit language knowledge, the one that is often inaccurate and effortful to access as it requires attention. In adopting a strong stance, Krashen (1981, 1985) distinguished (implicit) acquisition from (explicit) learning⁴, in which the former refers to the natural and subconscious L2 acquisition that results in fluent language competence, whereas the latter refers to the formal and conscious process of studying an L2 that results in knowledge available nearly only for

⁴ Although Krashen (1981, 1985) distinguished L2 ‘learning’ from ‘acquisition’, the present study treats both terms as synonyms to refer to general L2 development.

error correction. However, there is no agreement on whether L2 learning is implicit or explicit, resulting in implicit, explicit, or both types of knowledge, and whether they are related or not. Following that, three positions were raised on the *interface* of implicit and explicit knowledge, which will be presented in the next paragraphs based on the accounts of Ellis (2005) and Kim (2020).

The non-interface position views implicit and explicit knowledge as two distinctive systems that do not interact with each other. Proponents of this view claim that implicit and explicit knowledge are acquired through different processes, stored in different parts of the brain, and retrieved differently. In this way, one type of knowledge does not transform into or influence the other. This position is strongly advocated by Krashen's (1981, 1985) distinction between (implicit) acquisition and (explicit) learning. Krashen (1981, 1985) argued that implicit knowledge is acquired through natural, meaningful communication when the L2 learner is focused on meaning, that is, on understanding and being understood, whereas explicit knowledge is believed to be learned through focus on form (e.g., grammar), e.g., during traditional language lessons. As a result, explicit knowledge is restricted mostly to monitoring the language output and does not become implicit knowledge, which is the most important language competence for fluent language performance (Krashen, 1981, 1985).

The strong-interface position views implicit and explicit knowledge as two connected systems. Through practice, explicit knowledge can become implicit, and implicit knowledge can become explicit. This position is strongly associated with the Skill Acquisition Theory (Anderson, 1982), which focuses on how explicit language knowledge becomes automatized with frequent, conscious practice. Therefore, the initial acquisition of explicit knowledge, through language instruction, for instance, is a fundamental step to the later automatization of this explicit knowledge, driven by deliberate practice (DeKeyser, 2007). In accordance with the non-interface position, DeKeyser (2017) suggested that explicit knowledge does not simply convert into implicit knowledge, but rather, the repeated use of one system leads to the gradual establishment of the other, without diminishing the levels of the former.

The weak-interface position views implicit and explicit knowledge as two distinctive systems that are related to a limited extent. That is, explicit knowledge can influence the acquisition of implicit knowledge and vice versa, although they both remain distinct systems that are the products of different language educational experiences. This position is strongly supported by Rod Ellis (1994), Nick Ellis (2006), and Schmidt's (1990, 1995) notion of *noticing the form* and *noticing the gap*. *Noticing the form* is the process of becoming aware, through focal attention, of a linguistic aspect that has so far been ignored, whereas *noticing the gap* is

the ability to notice pronunciation deviations in one's own pronunciation in relation to the target form. According to Schmidt (1995), these are the two possible ways to become aware of linguistic aspects. Building on that, R. Ellis (1994) suggested that explicit knowledge assists the initial and final stages of L2 learning by facilitating the processes of *noticing the form* and *noticing the gap* and the conscious practice of novel linguistic forms. Explicit knowledge is, therefore, especially relevant in L2 acquisition as N. Ellis (2006) noted that although L1 acquisition may primarily involve implicit knowledge, L2 acquisition differs in terms of L1 transfer and learnt attention (i.e., where to focus attention). In this view, conscious explicit knowledge enables L2 learners to practice the regular linguistic forms acquired explicitly to the point of eventual automatization of the knowledge. Therefore, it is suggested that with practice, explicit knowledge can become automatized, that is, unverbalizable and effortlessly retrieved, to an extent when it is difficult to differentiate automatized explicit knowledge from implicit knowledge (Ellis, 2004).

From the three different positions on the relation between explicit and implicit knowledge, the weak-interface position seems the most plausible for two reasons. First, proponents of the three positions recognize that explicit knowledge interacts with implicit knowledge, although the degree of this interaction may vary. Second, no interaction between the two types of language knowledge seems implausible, as older children have been shown to develop L1 metalinguistic knowledge without necessarily receiving explicit instruction (Ellis, 2006). Second, the strong interaction also seems implausible because if that were the case, retrieval would be slow and inefficient due to extensive parallel processing (Kivistö-de Souza, 2015). However, more studies are needed to obtain a better understanding of the interface issue.

Despite the heated theoretical discussion on the interface positions, few studies have empirically investigated the link between explicit and implicit L2 knowledge. Yet, studies in L2 morphosyntax have done so (Kim; Godfroid, 2023; Suzuki; DeKeyser, 2017) and findings indicate an association between the two types of knowledge. The two studies will be further reviewed in the next two paragraphs.

Suzuki and DeKeyser (2017) were pioneers in empirically exploring the relationship between implicit and explicit L2 knowledge outside laboratory settings, in a naturalistic L2 setting. The cross-sectional study examined the influence of automatized explicit knowledge on implicit knowledge, along with the contributions of explicit and implicit learning aptitudes and phonological short-term memory on the acquisition of those two types of knowledge. More specifically, implicit and explicit knowledge about three Japanese linguistic structures was tested: (i) transitive/intransitive verb pairs, (ii) classifiers, and (iii) locative particles. The

participants were one hundred advanced learners of Japanese whose first language was Chinese, currently living in Japan for at least 2 years at the time. Three real-time⁵ comprehension measures focused on meaning were used to assess implicit knowledge, whereas three timed tasks focused on form were employed to assess automatized explicit knowledge. To test the interface issue, two structural equation models were run, which differed in terms of an additional path from automatized explicit knowledge to implicit knowledge in the ‘interface’ model that was not included in the ‘non-interface’ model. Although there was no statistically significant difference between the two models, the path from automatized explicit knowledge to implicit knowledge reached statistical significance in the interface model. Moreover, the paths from the three cognitive aptitudes to the two types of knowledge were not significant, except for the path from explicit learning aptitude to automatized explicit knowledge. Therefore, their results suggest that explicit learning aptitude predicted the acquisition of automatized explicit knowledge, which, in turn, had an impact on the acquisition of implicit knowledge, even though the test did not establish causality between the statistical relationship between automatized explicit knowledge and implicit knowledge.

Following that, Kim and Godfroid (2023) examined the longitudinal associations between explicit and implicit L2 knowledge and meaning- and form-focused activity types. Three tests of implicit knowledge and two tests of explicit knowledge targeting grammatical features were applied at the beginning and the end of a university semester. To assess engagement in activity types, five self-reported language exposure logs were administered throughout the semester. Two sets of cross-lagged path models were run to analyze the causal relationship between explicit and implicit knowledge over time. The participants were 149 advanced learners of English (122 participants remained in the second data collection) studying at a university in the American Midwest. Results revealed three main findings. First, the strongest predictor of current implicit knowledge was previous implicit knowledge, and the strongest predictor of current explicit knowledge was previous explicit knowledge. Second, prior explicit knowledge had a positive effect on the development of implicit knowledge, and prior implicit knowledge had a positive effect on the development of explicit knowledge. Third, form- and meaning-focused activity types did not predict either knowledge type development. Therefore, findings indicate a reciprocal association between implicit and explicit L2

⁵ ‘Real-time’ is frequently used as a synonym of ‘online’ tests to refer to measures of performance during the test execution, rather than after the test completion. In this case (Suzuki; DeKeyser, 2017), the real-time measurements used reaction time and eye-movement to assess implicit knowledge.

morphosyntactic knowledge. Both types of knowledge mutually and bidirectionally influenced each other, while activity type did not predict knowledge development.

Apart from the above studies examining the relationship between implicit and explicit L2 morphosyntactic knowledge, to the best of my knowledge, no studies have examined the relationship in other language domains, let alone L2 phonology. Ellis (2005) noted that one significant obstacle to empirically investigating the interface hypotheses is the lack of instruments agreed upon to measure each type of knowledge, as they overlap in performance. Following that, Roehr-Brackin (2024), acknowledging the importance of empirical studies on the interface hypotheses, proposed a research agenda in the field that highlights the need for the identification of valid and reliable measures of implicit and explicit L2 knowledge.

2.1.3 L2 phonological awareness

Despite the few instruments available in the literature to measure L2 morphosyntactic knowledge that one could consider as models for other language domains, the obstacle to creating measures of L2 phonological awareness is prominent given its unique aspects. Kivistö-de Souza (2015) mentioned three important considerations for creating instruments to measure L2 phonological awareness. The first consideration regards the visual presentation of the stimuli. As judgments of segment accuracy depend on the context (e.g., the vowel [æ] is different in ‘cat’ and ‘bad’), the stimulus needs to be presented visually in its context rather than solely as an individual sound. However, phonetically naïve language learners are unfamiliar with phonetic symbols, so the visual representation of the stimuli must be written orthographically, with letters. The second consideration involves the nature of individual sounds that are very short, often rapidly pronounced, and rarely encountered in isolation. This leads to a trade-off between presenting the phone accompanied by others in a real context, but risking losing its prominence, or presenting the phone in isolation with maximum salience, albeit somewhat artificial. The third consideration comes with the motoric limitations when pronouncing L2 sounds. Pronunciation deviations may not be a result of low levels of awareness but of such limitations. Therefore, oral production should not be regarded as an indication of L2 phonological awareness. These considerations together make the development of measures of L2 phonological awareness particularly difficult, and instruments for morphosyntactic L2 knowledge are not coherent models in this regard.

L2 phonological awareness consists of implicit and explicit knowledge about the phonological system of the target language. It varies in levels among the three language domains: segmental (individual phones), suprasegmental (e.g., intonation), and phonotactic (e.g., possible sound combinations) (Kivistö-de Souza, 2015). Kivistö-de Souza (2015) suggested that L2 phonological awareness consists mostly of implicit knowledge, given the inherently unconscious nature of speech, the lack of explicit instruction targeting L2 pronunciation, and the difficulty language users have in describing and explaining speech phenomena.

Phonological awareness has been extensively studied in L1 acquisition, targeting mainly children's literacy acquisition (Kivistö-de Souza, 2015). In this regard, L1 phonological awareness is generally operationalized as the ability to reflect upon and manipulate sounds, which entails mainly explicit knowledge. For instance, Oakhill and Kyle (2000, p. 152) defined phonological awareness as "the ability to detect, distinguish between, and manipulate the constituent sounds of words: syllables, onsets, rimes, and phonemes". Following that, the ability to replace phonemes to create new words, such as replacing the [b] in 'robe' with [p] to create 'rope', is one demonstration of phonological awareness. Research shows that higher levels of phonological awareness positively contribute to children's literacy acquisition (Ehri; O'Leary, 2022). Although studies in L1 phonological awareness concerning L1 literacy usually target explicit knowledge, implicit knowledge has also been acknowledged as part of L1 phonological awareness (e.g., Cunningham; Carroll, 2015; Geudens, 2006).

Alves (2012) brings the two constructs widely mentioned in studies of L1 phonological awareness, 'reflection' and 'manipulation', to the field of L2 acquisition. Reflection is regarded as the learner's ability to consciously reflect on phonology rather than using language as a mere method of communication. Manipulation is referred to as the ability to create new words and meanings by manipulating sounds (Alves, 2012). Despite the parallel between L1 and L2 phonological awareness, in a subsequent study, Alves (2024) highlighted the nature of L2 phonological awareness, which was argued to encompass two distinct but strongly related types of knowledge. The first type of knowledge is related to the specific phonological features of the L2, whereas the second type of knowledge involves the differences found between the L1 and the L2 phonological systems. Therefore, L2 phonological awareness is inherently different from L1 phonological awareness as the former entails additional knowledge to differentiate the sounds of the new language from the already well-established sounds of the L1. Alves (2024) suggested that possessing these two types of knowledge can benefit the perception and

production of L2 sounds. In other words, L2 phonological awareness would contribute to L2 perception and production.

Although L2 phonological awareness is considered a new field of interest within L2 speech learning (Kivistö-de Souza, 2021), research has found some relation between L2 phonological awareness and L2 perception and production (e.g., Baker; Trofimovich, 2006; Kennedy; Trofimovich, 2010; Kivistö-de Souza, 2015, 2017). The next paragraphs will present three important studies that have analyzed L2 phonological awareness. The first study presented (Kivistö-de Souza, 2017) is relevant here as it analyzed Brazilians' implicit knowledge about English phones and included VOT, which is analyzed in the present study. The second study (Flege; MacKay, 2004) employed a test (Experiment 4) used as the reference for the test measuring implicit awareness in the present study. Lastly, the third study (Kennedy; Trofimovich, 2010) analyzed L2 phonological awareness and L2 production and is widely cited in the field of L2 phonological awareness.

Kivistö-de Souza (2017) examined the implicit phonological awareness of English segments that Brazilian learners of English are shown to have difficulties with. The English segmental features included VOT, vowels, consonants, and final devoicing. To accomplish that, proficient Brazilian learners of English and American English speakers were asked to perform a perception test. The perception test consisted of native and non-native pronunciations of the target sounds, in which accuracy in identifying pronunciation deviations (non-native tokens) was regarded as participants' implicit awareness of English segments. For each trial, participants were shown the written word with the target sound underlined, but listened to the target sound only, followed by or preceded by another sound, and not to the entire spoken word. Results revealed that identifying deviations was difficult for Brazilian and American participants, whereas identifying native-like pronunciation as correct was not difficult. However, Brazilian learners scored poorer than Americans in both trial types. Concerning the pronunciation deviation types across segments, Brazilian learners' identification accuracy was above chance level only for the *Consonant* category, showing that Brazilian learners of English possess an overall low level of implicit awareness about English segments. More relevantly for the present study, *VOT* was found to be the most difficult category for both participant groups.

To examine implicit phonological awareness, the present study employs a perception test similar to the aforementioned one (Kivistö-de Souza, 2017), which resembles the *Error Detection Test* employed by Flege and MacKay (2004). Although Flege and MacKay (2004) do not state that they are measuring phonological awareness with the *Error Detection Test* (experiment 4), this appears to be the underlying object of investigation, considering previous

studies analyzing L2 phonological awareness through L2 perception only (e.g., Baker; Trofimovich, 2006; Shoemaker, 2014). Flege and MacKay (2004) analyzed the perception of English vowel pairs by Italian immigrants to Canada, differing in age of arrival and language use. The study consisted of 4 experiments, and in Experiment 4, it was tested if Italians who arrived in Canada at early stages of life and differed in their amount of Italian use (early-high and early-low) would perform differently from native English speakers in a more difficult perception test, given that both Italian groups and native speakers performed similarly in Experiment 3 (categorical discrimination). Therefore, Experiment 4 consisted of an *Error detection test* in which participants listened to short English phrases containing accurate and inaccurate samples of the English high front vowels /i/ and /ɪ/ and had to decide if the samples were correct or incorrect. The accurate stimuli were produced by native speakers and correctly transcribed by a set of listeners, and the same was done to the inaccurate stimuli produced by native Italians. In the test, participants were presented with the written phrase they were listening to. In each phrase, the target sound was replaced by an asterisk to ensure the target vowel to be judged was evident. Results showed that the early-high group (high L1 use) presented significantly lower scores than native English speakers, while the early-low group (low L1 use) did not differ significantly from native speakers. This suggests that L1 use may influence the perception of L2 vowels. Therefore, if Experiment 4 is considered to measure implicit L2 knowledge, it could be argued that language exposure and use (the proportion of L1 and L2 exposure and use frequency, in their case, Italian and English) affect implicit phonological awareness.

Kennedy and Trofimovich (2010), apart from finding a positive relationship between L2 awareness and production, also found a positive relationship between L2 qualitative awareness and L2 exposure and use. Ten language learners were asked to write weekly dialogue journals and self-rate their L2 daily use, among other tests. It was observed that those learners who received better pronunciation ratings were also the ones with more qualitative awareness comments, that is, viewing language as a means of communication rather than a set of items. Moreover, those learners who wrote more qualitative comments also reported more L2 listening. This finding suggests that L2 awareness may be related to L2 exposure and use.

L2 use can be said to likely matter to L2 phonological awareness, given that language awareness is thought to develop through the mechanisms of *noticing* (Schmidt, 1995). As discussed in the weak-interface position, the most plausible one, Schmidt's idea is that *noticing the form* and *noticing the gap* are the two processes by which one can become aware of linguistic aspects. Following that, the more contact the language learner has with the target

language, the more opportunities are available for the so far ignored linguistic aspect to be noticed. However, when some linguistic aspects seem hard to notice through L2 use, consciousness-raising activities and explicit instruction may be helpful (Teló; Kivistö-de Souza, 2023).

The next section will discuss the role played by L2 use in language awareness and in language acquisition as a whole.

2.1.4 L2 input quantity and quality

As mentioned above, L2 use likely matters to L2 phonological awareness. Indeed, studies have shown that the amount and nature of contact with the target language are significant predictors of successful L2 acquisition, especially for L2 phonology (Flege; Bohn, 2021). Therefore, L2 use may be better framed as *L2 input quantity and quality* (Moyer, 2009). *L2 input* can be defined as “all L2 vocal utterances the learner has heard and comprehended, including his own, regardless of whether these utterances have been produced correctly by L2 native speakers, or incorrectly by other non-native speakers of the L2” (Flege, 2009, p. 175). Moreover, the supply of L2 input in terms of quantity and quality differs greatly between naturalistic and instructed language learning settings (Muñoz, 2009).

In studies within naturalistic language learning settings, which mostly analyze participants who are immigrants, L2 input *quantity* is frequently operationalized as length of residence and proportion of L1 and L2 use (Mayor, 2009). Evidence from studies analyzing participants differing in the time spent in a country where the target language is spoken suggests that the quantity of L2 input plays a role in L2 acquisition (e.g., Kupske, 2016). In this way, language learners seem to benefit from longer lengths of residence, which is partly explained by greater amounts of L2 input (Moyer, 2009). However, that is not always the case, and some language learners show highly accented L2 speech even after spending years abroad (Flege, 2009). This may be partially explained by the proportion of L1 and L2 use, as just living abroad does not guarantee contact and use of the target language. Immigrants differ in their attitudes and beliefs towards the new language and country so that immigrants who are more integrated into the new community may also use the new language more frequently and this proportion of L1 and L2 use together with LOR may be a better indication of quantity of L2 input in this language learning setting (Flege; Bohn, 2021; Kupske; Lima Jr., 2022; Kupske; Lima Jr.; Alves, 2023).

In instructed language learning settings, L2 input is frequently operationalized as years of instruction, time on task (e.g., time spent studying the language), and proportion of L1 and L2 use (Mayor, 2009). Although years of instruction and time on task may consist of more straightforward measures as they are closely related to formal language learning such as taking language classes, studies in this language learning setting have also made an effort to include more complex estimates of proportion of L1 and L2 use in order to encompass the language input received outside the traditional language classroom giving the way technology has expanded the possible means of contact with foreign languages (Otto, 2017). Teló and Kivistö-de Souza (2023), for example, operationalized L2 input quantity as ‘L2 experience’ and ‘L2 use’ (based on self-estimates obtained through questionnaires) in which the former consisted of the number of years receiving formal language education and the latter consisted of “the percentage of time [...] spent engaging in oral communication in English over the past five days, five weeks, five months, and five years” (Teló; Kivistö-de Souza, 2023, p. 205). In their study (Teló; Kivistö-de Souza, 2023), the quantity of L2 input was not a significant predictor of English segmental production accuracy among Brazilian learners, whereas the quality of L2 input was. Following that, Saito and Hanzawa (2018) did not find L2 input quantity to predict segmental accuracy in Japanese learners, but prosodic accuracy (words and intonation). However, due to the low number of studies, the role of L2 input on foreign language acquisition in instructed settings remains inconclusive.

As suggested by the findings from Teló and Kivistö-de Souza (2023) and Saito and Hanzawa (2018), L2 input *quality* also matters for L2 speech acquisition. In such studies where L2 input quantity seems not to be relevant for L2 speech acquisition, this apparent limited contribution of L2 input may be better analyzed and interpreted if considering not only the amount but also the nature of contact with the L2, in other words, L2 input *quality* (Flege; Bohn, 2021). Thereupon, high-quality L2 input, that is, not only contact with input by native speakers, but also complex opportunities for meaningful communication using the L2, may affect the acquisition of L2 sounds (Flege, 2009; Flege; Bohn, 2021; Mayor, 2009).

Despite its importance, the role of L2 input quality on L2 phonetic acquisition has often been neglected, which may be partially due to measurement limitations (Flege, 2009; Moyer, 2009). Both in naturalistic and instructed language learning settings, L2 input quality is generally operationalized as the amount and type of interactions with native speakers. These interactions are usually assessed through self-reports on questionnaires, such as percentage or rating scales (e.g., Teló; Kivistö-de Souza, 2023). Such self-estimations are limited to the participants’ ability to provide accurate retrospective information (Flege, 2009), which is

likewise the case for measurements of L2 input quantity that rely on retroactive self-estimates, e.g., years of instruction and proportion of L1 and L2 use. Therefore, results may vary according to the measure and operationalization of the construct.

Considering the relevance that L2 input has shown to the development of L2 domains (Moyer, 2009), it appears pertinent to analyze the role of L2 input on L2 phonological awareness. However, only a handful of studies to date have ventured into this analysis, and findings are still inconclusive. While Venkatagiri and Levis (2007) did not find a relation between the amount of L2 exposure and the participants' explicit L2 phonological awareness, Shoemaker (2014) identified a relation between language experience and implicit aspects of L2 phonological awareness, and Kennedy (2012) found patterns in L2 use and L2 qualitative awareness at the individual level but not at the group level.

Venkatagiri and Levis (2007) investigated the correlations of explicit L2 phonological awareness and L2 speech comprehensibility along with other individual variables, including L2 exposure. L2 exposure was regarded as years of L2 instruction and months living in the L2 country. Thus, L2 input quality was not included in the analysis. No significant correlations were found between L2 exposure and L2 phonological awareness or L2 speech comprehensibility.

Shoemaker (2014) examined L2 learners' perception of allophonic variation as a word boundary cue, which Kivistö-de Souza (2015) attributed to implicit L2 awareness. L2 exposure was interpreted as years of instruction, thus also giving precedence to the quantity of L2 input over quality. Results demonstrated that third-year students were significantly more sensitive to allophonic variation signaling word boundaries than first-year students. This finding supports the idea that L2 input matters to the development of L2 speech, at least when it comes to L2 perception.

In light of previous findings indicating that L2 quantitative awareness is related to L2 use (Kennedy; Trofimovich, 2010), Kennedy (2012) sought to clarify the relationship between quantitative and qualitative L2 awareness and L2 use through more fine-grained data on L2 use. For that, ten language learners were asked to record their L2 use across ten weeks on a Language Activity Log. This log encompassed six main domains for language use, including formal and informal domains of social interaction, and activities involving mostly speaking and listening. Although this instrument comprised the type, number, and length of activities, it did not include the quality of learners' engagement. Moreover, quantitative and qualitative language awareness were measured through journal entries. Results yielded no significant association between any measures at the group level. However, three longitudinal patterns of the association between

awareness and L2 use were evident at the individual level. The first patterns revealed an increase in qualitative awareness without a concurrent increase in L2 use, which was attributed in part to the learner's high levels of qualitative awareness displayed since the beginning of testing. Thereupon, it could be argued that increased qualitative awareness contributes to even higher awareness without increased L2 use. The second and third patterns revealed positive associations between either increased qualitative awareness corresponded to increased L2 use, or decreased qualitative awareness coincided with decreased L2 use. This finding suggests that at least for some learners, a self-reinforcing reciprocal association between L2 awareness and L2 use exists.

Therefore, considering the studies reviewed, the impact of L2 input quantity and quality on L2 phonological awareness remains unclear. The mixed empirical evidence might be due, at least in part, to the limitations of the current methods employed to measure L2 input quantity and quality. Mayor (2009) calls for more sophisticated measurements of L2 input quantity and quality that are not completely reliant on learners' self-reports on quantitative scales (e.g., Likert scales). Although still employing such methodological tools, the present study aims to contribute to the growing body of research on the role played by L2 input in implicit L2 phonological awareness.

After discussing the broad topic of language awareness, the next section is dedicated to VOT, the phonetic feature analyzed in the present study.

2.2 VOICE ONSET TIME (VOT)

Voice Onset Time (VOT) is a phonetic-phonological feature that occurs in the production of stop consonants. Consonants are speech sounds produced with some contact between articulators (e.g., lips and tongue) in the vocal tract, with partial or total air passage obstruction (Yavas, 2011). The nature of the contact between articulators is classified with most regard to manner and place of articulation, which is how, along with voicing, consonants are generally classified. Regarding the manner of articulation, consonants can be classified as stops, fricatives, and affricates, for example. Stops, like [p, b, t, d, k, g], also known as plosives, are the consonants produced with a total blockage of the airflow by the closure of two articulators followed by the release of the closure, generating, then, a 'burst' (plosion) (Kupske, 2016). With respect to place of articulation, stop consonants, the ones in focus here, can be classified

as bilabial, alveolar, and velar, for example⁶. Bilabial stops [p, b] result from the closure of both lips. Alveolar stops [t, d] are created by approaching the tongue tip or blade to the alveolar ridge⁷ (the bumpy area behind the upper front teeth). Velar stops [k, g] are produced in the obstruction between the back of the tongue and the soft palate (Yavas, 2011). Both English and BP phonologies contrast these three sets of stops in minimal pairs. In other words, the meaning of words is affected by the replacement of a bilabial stop with an alveolar or velar stop, e.g., English: ‘pan–tan–can’ [pæn]–[tæn]–[kæn]; BP: ‘pão–tão–cão’ [pẽw]–[tẽw]–[kẽw].

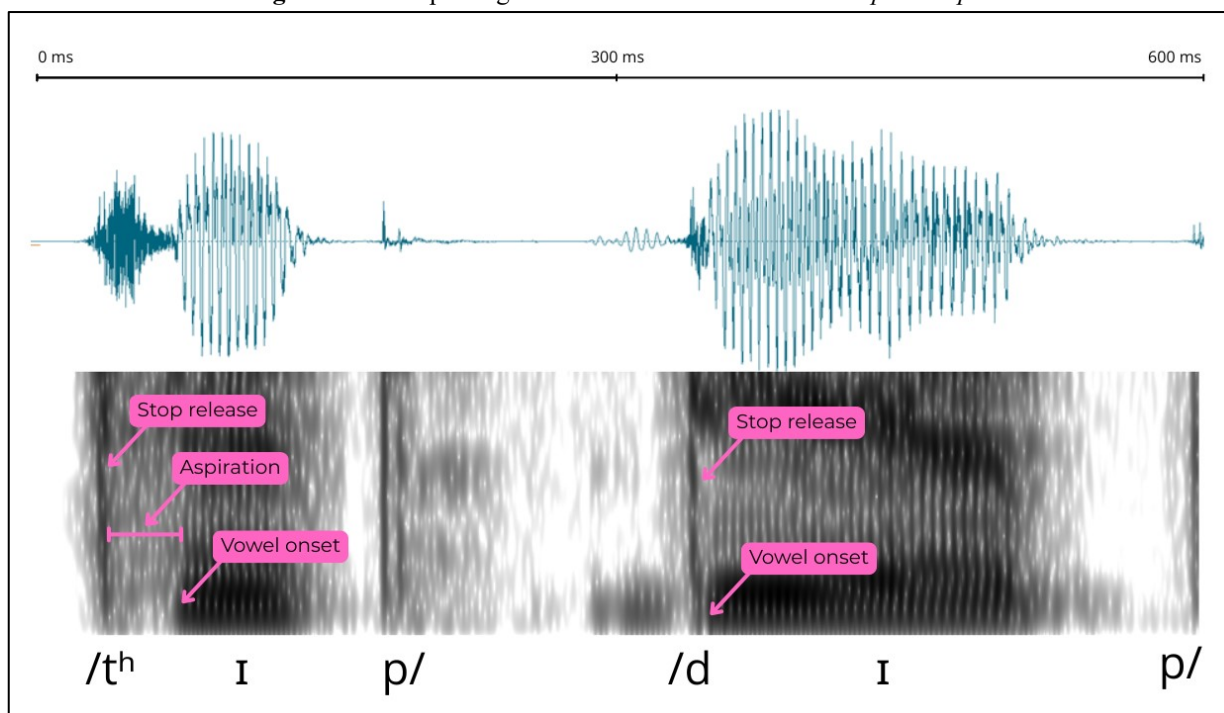
Along with manner and place of articulation, consonants can also be classified regarding voicing, that is, the existence or not of vocal fold vibration. Voiced stops [b, d, g] present vocal fold vibration during the articulatory closure, while voiceless stops [p, t, k] do not (Lisker; Abramson, 1964). Therefore, voicing can be used to differentiate consonants falling within the same articulation category, such as the bilabial stops [b, p]. In BP, examples of minimal pairs differing in the voicing of stop consonants are: pato–bato [‘pa.tu]–[‘ba.tu], tela–dela [‘tɛ.lɐ]–[‘dɛ.lɐ], and cama–gama [‘kẽ.mɐ]–[‘gẽ.mɐ]. However, the voiced/voiceless distinction is not entirely suitable for some languages, such as most varieties of English. As Lisker and Abramson (1964) noted, in English, [b, d, g] are fully voiced only in medial position (i.e., in between vowels), but when pronounced in initial position, those stops regarded as voiced tend to be pronounced with silent closure intervals, without vocal fold vibration, and consequently they could be considered voiceless.

As a solution to separate the two sets of English stops, the aspiration feature can be used in addition to voicing, given that English voiceless stops exhibit aspiration when pronounced in initial position, whereas voiced stops do not (Lisker; Abramson, 1964). Aspiration is the voiceless gap between the stop release and the start of voicing of the following vowel, and the distinction between aspirated and unaspirated sounds is based on the time difference between this voiceless gap (Ladefoged; Johnson, 2011). In English, examples of minimal pairs differing in VOT are: ‘pack–back’ [pæk]–[bæk], ‘tip–dip’ [tɪp]–[dɪp], ‘came–game’ [keɪm]–[geɪm]. In phonetic transcriptions, aspiration is represented by a small, raised *h* (^h), and in acoustic representations (waveforms and spectrograms), it can be seen as a gap with small amplitude presenting semi-random periodicity right after the stop release (Figure 1) (Ladefoged; Johnson, 2011).

⁶ In American and British English, stop consonants can also be classified as glottal, which “occurs when the vocal cords are held tightly together” (Yavas, 2011, p. 61), and is frequently found as an allophone of /t/ in syllable-final position, such as in ‘setback’ [‘sɛʔ.bæk] (Yavas, 2011).

⁷ In BP, the tongue also touches the top of the front teeth (Kupske, 2016).

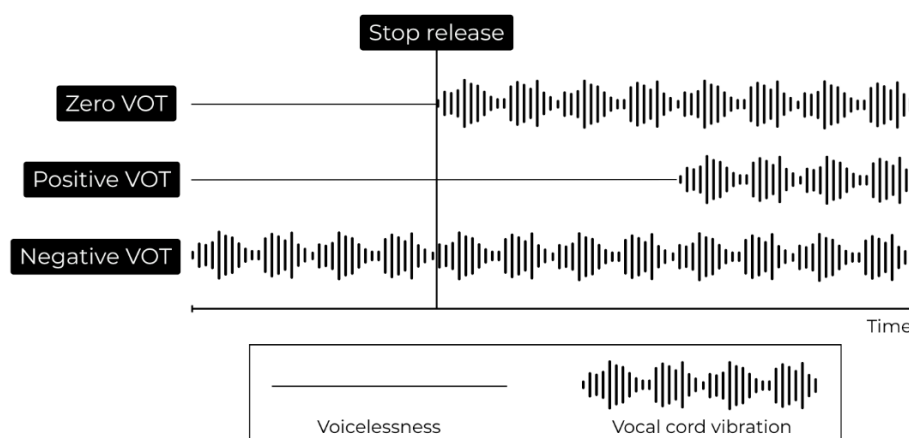
Figure 1 – The spectrogram and waveform of the words *tip* and *dip*



Source: the author.

By measuring the aspiration of a stop in milliseconds (ms), like in the spectrogram shown in Figure 1, its voice onset time (VOT) is what is being measured. VOT is the time difference between the stop release and the beginning of vocal fold vibration (before or after the stop release) (Yavas, 2011). Note that aspiration is the voiceless gap between the stop release and the start of voicing, while VOT is the time difference between them. Therefore, VOT is useful to measure and categorize not only voiceless stops but also stops with no aspiration. To that end, Lisker and Abramson (1964) established a way to measure VOT value by setting the moment of stop release as Zero, so negative values are used when voicing starts before the release, and positive values when voicing starts after the release.

Stop consonants can be classified into three VOT categories: *Negative VOT*, *Zero VOT*, and *Positive VOT* (Lisker; Abramson, 1964). *Negative VOT* (also known as *pre-voicing* or *voicing lead*) happens when the vocal fold vibration starts before the stop release. *Zero VOT* (*short-lag*) refers to the little time difference between the beginning of vocal fold vibration and the stop release. *Positive VOT* (*aspiration* or *long-lag*) occurs when aspiration is created by the delay of vocal fold vibration after the stop release. Therefore, different languages may present different VOT patterns, as with English and BP. Figure 2 illustrates the three VOT categories. The stop release is represented in a vertical line, and time in a horizontal line. Similarly to waveforms, voicelessness is represented by a straight line, while voicing is represented by a bar-style sound wave.

Figure 2 – Voice Onset Time categories

Source: Adapted from Cohen (2004, p. 5)

The mean VOT values in milliseconds (ms) for each category presented by Lisker and Abramson (1964) are: -125 to -75 ms for Negative VOT, from 0 to +25 ms for Zero VOT, and from about +60 to +100 ms for Positive VOT. Although studies like Lisker and Abramson (1964) have tried to find mean VOT values for the stops of different languages, findings have suggested that this phonological feature is not static but is influenced by many aspects. One major effect on VOT value is the quality of the following vowel (França, 2011; Klatt, 1975; Yavas, 2008). In this regard, it appears that higher following vowels result in longer VOT values. This also seems true for stressed syllables (Yavas, 2011) and slower speech rate (Miller; Volaitis, 1989). In addition, Ryalls *et al.* (1997) and Whiteside and Irving (1998) found that women generally pronounce longer VOT than men; however, Allen *et al.* (2003) argued that such a gender difference is not seen when speech rate is controlled among speakers. Thereupon, Prestes (2013) called attention to the necessity of carefully considering studies examining VOT as values can vary extensively depending on how the data was collected and analyzed. Altogether, those studies presenting mean VOT values can still be helpful to research languages, mainly if considering the singularities of each study. The next subsections will discuss English and BP VOT, bringing VOT mean values to exemplify and compare the two language phonologies.

2.2.1 English VOT

In English, VOT is the most important phonetic cue for distinguishing voiceless from voiced stops (Alves; Kampff, 2019). Therefore, native English speakers heavily depend on VOT to contrast minimal pairs like *pet-bet*, *tip-dip*, and *coat-goat*. In word-initial position,

English voiceless stops [p, t, k] are generally produced with *Positive VOT*, and English voiced stops [b, d, g] with *Zero VOT* (yet productions with *Negative VOT* are possible) (Schwartzhaupt; Alves; Fontes, 2015). Given the importance of VOT, many studies have tried to obtain mean VOT values for the English stops in different contexts. The following paragraphs will report on some of the most addressed studies. Table 2 summarizes the information.

Lisker and Abramson (1964), one of the pioneering studies analyzing VOT, found the following VOT mean values for the English voiceless stops: 58 ms for [p], 70 ms for [t], and 80 ms for [k]; and the following values for the voiced stops: 1/-101 ms for /b/, 5/-102 ms for [d], and 21/-88 ms for [g]. The authors (Lisker; Abramson, 1964) provided two sets of values for the voiced stops, not to lump positive and negative values. The study analyzed four speakers of American English, and more information on the subjects and context is not provided.

Table 2 – Mean VOT values in ms for the voiceless and voiced stops in English

Authors	Subjects	English stops						Context
		Voiceless			Voiced			
		/p/	/t/	/k/	/b/	/d/	/g/	
Lisker; Abramson (1964)	4	58	70	80	1/-101	5/-102	21/-88	-
Baran <i>et al.</i> (1977)	3 F	50.7	46.1	49.6	0.7	8.9	18	Adult-directed conversation
		35.7	37.6	41.4	-0.3	10.3	16.2	Child-directed conversation
		64.3	70.6	79.2	-16.4	-5	-4.5	Single-word citation form
Kent; Read (2002)	134	46/85	65/95	70/110	1/20	0/21	14/35	Maximum and minimum means across 12 studies
Chodroff <i>et al.</i> (2015)	69 F 60 M	50.8	60.5	54.4	8.7	13.8	17.2	Sentences
Kupske (2016)	5 F 5 M	56.95	77.31	82.55	-	-	-	Carrier sentences
Chodroff; Wilson (2017)	13 F 11 M	89	98	99	13	21	28	Isolated speech
	102 F 78 M	51	61	56	8	14	17	Connected speech

Source: the author.

Key: F = female; M = male.

Baran et al. (1977) analyzed the VOT of three adult female speakers in three different conditions: (i) adult-directed conversation, (ii) child-directed conversation, and (iii) recitation of citation forms. Results showed significant differences between citation forms and conversational speech, but not between adult-directed and child-directed conversation. When

comparing the VOT mean values found in the study (Table 2), it is possible to see that English voiceless stops were produced with longer VOT values in citation form than in conversation. Conversely, English voiced stops were produced with negative values only in citation form but not in conversation, except for [b] in the child-directed conversation (-0.3 ms). Findings are in line with previous studies that found English voiceless stops to have shorter VOT values in connected speech than in readings (Lisker; Abramson, 1967).

Kent and Read (2002) presented minimum and maximum VOT mean values from a compilation of twelve studies analyzing English VOT (Auzou *et al.*, 2000). Most of the studies analyzed stops in target words inserted in carrier sentences, e.g., ‘Say ____ again’. By summing up the twelve studies, there were 134 subjects. The maximum VOT mean value for [k] represents the maximum value from the studies included in this review and in Table 2, which may be due to the large body of studies compiled in Auzou *et al.* (2000). The wide range of values in Kent and Read (2002) evidences that VOT values are not categorical.

Chodroff *et al.* (2015) analyzed the VOT of 69 female and 60 male American English speakers, aged from 19 to 87 years old. The target stops were inserted in sentences to be read aloud. The mean VOT values are displayed in Table 2. Contrary to some studies (e.g., Lisker; Abramson, 1964), the mean VOT values of the English voiceless stops [p, t, k] did not receive ascending values. The value for [k] (54.4 ms) was lower than for [t] (60.5 ms). Such a departure from ascending patterns is also seen in Baran *et al.* (1977) and Chodroff and Wilson (2017).

Kupske (2016) analyzed the VOT of voiceless stops from ten Standard Southern British English speakers. Speakers were 5 females and 5 males with a mean age of 24.55 years. The values align with Lisker and Abramson (1964), who analyzed American English speakers. The target stops were inserted in the initial position of monosyllabic words, followed by a high-low or a back-low vowel, which were inserted in the carrier sentence ‘I would say ____’.

Lastly, Chodroff and Wilson (2017) analyzed the VOT of American English voiceless and voiced stops in two different contexts: isolated and connected speech. For the isolated speech context, 13 females and 11 males pronounced stop-initial CVC syllables in the carrier sentence ‘Say _____ again’. Mean VOT values for voiceless stops in isolated speech were considerably high, which, according to the authors’ speculation, is a reflection of an overall slow speaking rate. For instance, the bilabial and alveolar [b] and [t] display the highest mean values from the studies in Table 2. For the connected speech context, recordings of speech sessions by 102 females and 78 males were analyzed, in which stops appeared word-initially and were followed by vowels. Mean VOT values for voiceless stops in connected speech were overall shorter than values for isolated speech, which is in line with Baran *et al.* (1977).

In the present study, the mean VOT values from Chodroff and Wilson (2017) in isolated speech will be used as the reference for English voiceless stops for three reasons. First, Chodroff and Wilson's (2017) values are the closest to the mean VOT values found in the production of the American native speakers of the present study (82.4 ms for /p/, 102.1 ms for /t/, and 112.8 ms for /k/). Another reason stems from the elicitation technique adopted by Chodroff and Wilson (2017), the carrier sentences, which is the same adopted in the present study. Additionally, Chodroff and Wilson (2017) is the most recent study encompassed in this review (Table 2).

2.2.2 Brazilian Portuguese VOT

Contrary to English, in which voiceless and voiced stops are better distinguished by VOT rather than voicing, in BP, VOT does not contrast these two sets of stops. Instead, VOT plays a marginal role in the identification of PB voiced and voiceless stops, and native speakers of BP use phonetic cues like burst intensity, *f0* (*fundamental frequency*, perceived as pitch by listeners, which refers to the rate of vocal fold vibration, thus, voicing), or even the combination of different phonetic cues, to distinguish [p, t, k] from [b, d, g] (Alves; Kampff, 2019; Kupske; Perozzo, 2024). In word-initial position, the PB voiceless stops are generally produced with *Zero VOT*, and voiced stops with *Negative VOT* (Schwartzhaupt; Alves; Fontes, 2015). However, some studies analyzing VOT productions from the Southern region of Brazil have shown that BP voiceless stops, especially the velar stop [k], can also be produced with *Positive VOT*, and that BP voiced stops can also be produced with *Zero VOT* (Alves *et al.*, 2008; Alves, Zimmer, 2015). These inconsistencies, at least to a smaller degree, are part of the argument that VOT is not the main acoustic cue used by Brazilians in distinguishing voiced from voiceless stops (Alves, Zimmer, 2015). Alternatively, voiced stops in BP, just like in English, are generally produced with a weaker burst intensity (Alves, Zimmer, 2015), which could account in part for that distinction by Brazilian speakers of BP as L1.

Although VOT is not the most important phonetic aspect for distinguishing voiceless from voiced stops in BP, many studies, mainly in SLA, have obtained mean VOT values for BP. The next paragraphs will report on their findings. Table 3 summarizes the information.

Istre (1980 *apud* Klein, 1999)⁸, one of the pioneering studies analyzing BP VOT, analyzed the VOT productions of four males from the state of Santa Catarina, in Southern

⁸ ISTRE, G. L. **Um estudo do VOT em monolíngües brasileiros**. Unpublished manuscript, Universidade Federal de Santa Catarina, Florianópolis, Brazil, 1980.

Brazil. The voiceless and voiced stops were inserted in monosyllabic CV (consonant + vowel) words, with the following vowels [i, e, ε, a, ɔ, o, u]. The voiceless stops [p, t, k] received ascending positive mean values of 11.95 ms, 18.48 ms, and 38.53 ms, respectively, whereas the voiced stops [b, d, g] received negative mean values of -39.45 ms, -60.59 ms, and -53.69 ms, respectively.

Klein (1999) obtained mean VOT values of two female and two male speakers of BP, aged between 20 and 40 years old. The target stops were inserted in two contexts: sentences and isolated words, followed by the vowels /e, a, o/. The following vowels were selected by their classification. Only close-mid front and back vowels and low central vowels were selected, rather than high front vowels, because the combination of [i] with the velar stops [t] and [d] in PB results, in most cases, in the affricates [tʃ] and [dʒ]. The mean VOT values for the voiceless stops in both contexts are comparable to Istre's (1980) values. The values for the voiceless stop [p] are the most divergent from Istre's (1980) value for the same stop. Moreover, the voiced stops received considerably lower mean values in comparison to Istre (1980), particularly the value for [b] in sentences.

Table 3 – Mean VOT values in ms for the voiceless and voiced stops in PB

Authors	Subjects	Brazilian Portuguese stops						Context
		Voiceless			Voiced			
		/p/	/t/	/k/	/b/	/d/	/g/	
Istre (1980)	4 M	11.95	18.48	38.53	-39.45	-60.59	-53.69	Monosyllabic segments (CV)
Klein (1999)	2 F	15.49	17.13	33.9	-104.98	-93.31	-76.93	Sentences
	2 M	17.70	16.78	37.82	-86.57	-99.2	-91.16	Isolated words
Alves <i>et al.</i> (2008)	35	37.49	40.69	47.24	-	-	-	Aspirated
		14.96	18.28	17.18	-	-	-	Unaspirated
Alves (2015)	5 F	21	20	55	-93	-87	-72	Sentences
Kupske (2016)	5 F 5 M	15.13	17.88	38.93	-	-	-	Carrier sentences

Source: the author.

Key: F = female; M = male.

Alves et al. (2008) analyzed the alleged aspiration of voiceless stops in BP. For that, semi-spontaneous speech productions of 35 speakers of BP, mainly from the South and Southeast of Brazil, were analyzed. Results were presented as the frequency of aspirated and unaspirated productions of each voiceless stop analyzed, with different sets of values for the aspirated and unaspirated productions. Findings indicated that BP velar [k] and alveolar [t] stops

were more frequently produced with aspiration, that is, with VOT at a minimum between 35 ms and 40 ms. The bilabial stop [b] was produced with aspiration in approximately 50% of the samples. The authors point out that BP voiceless stops seem to be classified in an intermediate region of aspiration, as slightly aspirated stops, contrary to English voiceless stops, which are evidently aspirated stops, in the long-lag category. Moreover, the position of stress in the syllable was shown to affect VOT values. Findings from Alves *et al.* (2008) evidence the presence of aspiration in the production of BP voiceless stops, although BP VOT values are still generally lower than the values for the English voiceless stops.

Alves (2015) analyzed the VOT productions of stops inserted in carrier sentences spoken by 5 female BP speakers from Santa Catarina. The speakers were around 21 and 26 years old. Apart from the mean VOT values of the aspirated productions in Alves *et al.* (2008), the VOT mean values for the BP voiceless stops in Alves (2015) represent the highest values among the studies included in Table 3. The mean value for the velar [k] (55 ms) was even higher than the mean for the aspirated [k] in Alves *et al.* (2008) (47.24 ms). Alves (2015, p. 150) addressed that the velar [k] was aspirated in 87% of the productions, in which 67% were slightly aspirated (values between 35-55 ms), and 33% aspirated (values between 55-95 ms). Findings agree with Alves *et al.* (2008) and show that the velar [k] can be classified as having *Positive VOT* in BP.

Lastly, Kupske (2016) also analyzed the VOT productions of stops inserted in carrier sentences by speakers of BP from Southern Brazil. Speakers were 5 females and 5 males born in Porto Alegre, Rio Grande do Sul, with a mean age of 26.9 years. The mean VOT values obtained for the voiceless stops are consistent with those from Istre (1980) and Klein (1999).

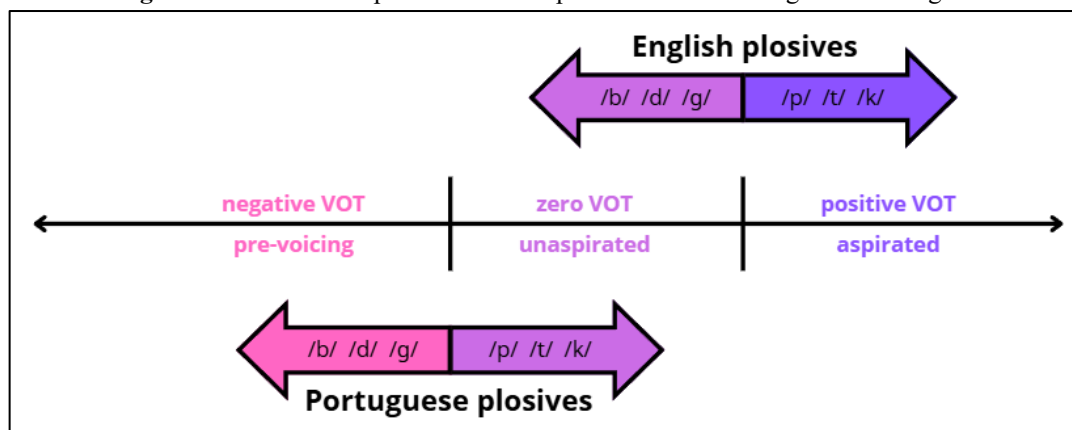
The selection of the mean VOT values used as reference for BP voiceless stops in the present study followed similar criteria to those for the selection of English mean VOT values. The mean values from Kupske (2016) will be used as a reference for BP voiceless stops. Kupske's (2016) values align with Istre's (1980) values, which have been used in other studies (e.g., Kupske; Oliveira, 2020; Kupske, Perozzo, 2024). But contrary to Istre (1980), Kupske (2016) consists of a much more recent study with more speakers of balanced gender. Moreover, the elicitation technique used was also carrier sentences, like the one employed in the present study.

2.2.3 Acquisition of English VOT by Brazilians

As VOT does not play a central role in the identification of stops in BP, native speakers of the language may not give the necessary attention to this important phonetic cue in English (Alves; Kampff, 2019). In fact, studies have shown that VOT is not the most relevant cue used for distinguishing English [p, t, k] from [b, d, g] by Brazilian learners of English, who were also shown to produce English stops with VOT values resembling the values of their L1 (Alves; Zimmer, 2015; Kupske; Oliveira, 2020), that is, their English voiceless stops are generally produced with shorter VOT values than those of native English speakers.

Although both English and BP comprise two VOT categories, *Zero* and *Positive VOT*, and *Negative* and *Zero VOT*, respectively, their voiced and voiceless stops fall within different categories (Figure 3). These differences in the stop consonants can cause difficulties for Brazilian learners of English who may not be aware of the English VOT patterns. Brazilians may not perceive such a difference between their L1 and L2, and BP VOT patterns may be transferred to the production of English VOT, which eventually affects L2 intelligibility (Alves; Kampff, 2019; Kupske; Oliveira, 2020). As an illustration, Brazilian learners producing English voiceless stops (*Positive VOT*) with VOT values resembling their L1 (*Zero VOT*) may be misunderstood by English speakers who use VOT to distinguish voicing, and thus perceive those English voiceless stops as voiced (*Zero VOT*), such as ‘pet’ being perceived as ‘bet’ (Alves; Zimmer, 2015). The next paragraphs will report on the major studies that have analyzed the perception and the production of English stops by Brazilian learners of English, some of which have also analyzed the effects of explicit instruction.

Figure 3 – Commonest production of stop consonants in Portuguese and English



Source: adapted from Falkievicz (2024, p. 38)

Schwartzhaupt, Alves, and Fontes (2015) analyzed the perception (discrimination and identification) of initial stops in English by 17 advanced Southern Brazilian learners of English and 20 native speakers of American English. The discrimination test consisted of an AxB test made of word triads in which participants were asked to indicate if the initial consonant of the second word was similar to that of the first word (AAB), to that of the second word (ABB), or if the three words had the same initial consonant (AAA). In the identification test, participants had to label the initial consonant of each production within six options: [p, b, t, d, k, g]. The stimuli consisted of four VOT patterns: Negative and Zero VOT from voiced stops, Positive VOT from voiceless stops, and one additional manipulated pattern, the Artificial Zero VOT. For that, the voiceless stops with Positive VOT had their VOT cut off, which hypothetically should sound like productions of voiced stops, as they present the VOT pattern typically found in voiced segments – Zero VOT. Results showed that Americans and Brazilians behaved similarly in their perception of Negative and Zero VOT as associated with voiced stops, and Positive VOT as associated with voiceless stops, while the difference between the two groups was found for the Artificial Zero VOT. Whereas the absence of positive VOT values was perceived as voiced stops by Americans, this new manipulated Artificial Zero VOT pattern was still taken as voiceless stops by Brazilians. This interesting finding suggests that cues other than VOT, such as burst intensity or the f_0 value of the following vowel, are more relevant in Brazilian learners' perception of the voiced and voiceless feature of English stops.

Alves and Zimmer (2015) analyzed the perception (identification and discrimination) and production of initial stops in English by 34 Southern Brazilian learners (24 beginners and 10 advanced learners). The study included identification and discrimination tests similar to those of Schwartzhaupt, Alves, and Fontes (2015), along with two production tests, one in English and one in BP, which consisted of reading words initiated with [p, b, t, d, k, g] and [p, b, k, g], respectively. Also employing the manipulated Artificial Zero VOT pattern in the perception tests, the study found that even advanced Brazilian learners did not use VOT as the main cue to distinguish [p, t, k] from [b, d, g], as both groups of learners tended to discriminate Negative VOT as distinct from Artificial Zero VOT – which hypothetically should sound similarly if VOT was taken as the main cue in perceiving stops as voiced or voiceless. In the production tests, Brazilians produced longer VOT values in English voiceless stops compared to their BP productions, although still shorter than VOT values by native speakers of English found in the literature. This led to the conclusion that Brazilians recognize that English voiceless stops are produced with longer VOT values than those in their L1, perhaps identifying English VOT as an “allophonic” variation that should be employed in their L2 speech to reduce

accent. However, Brazilian learners still did not perceive VOT as the main cue to distinguish voiced from voiceless stops. Alves and Zimmer (2015) suggested, then, that explicit instruction on VOT can help elucidate the phenomenon.

Regarding pronunciation instruction, Kupske and Oliveira (2020) found positive effects of explicit pronunciation instruction on English VOT production of voiceless stops by Brazilian learners of English. Kupske and Oliveira (2020) obtained English VOT values from 16 BP speakers in three data collections: a pre-test, an immediate post-test, and a delayed post-test. Brazilians were divided into control and experimental groups, in which the latter received a 2-hour instruction session on English stops after the pre-test based on the 5-step model proposed by Celce-Murcia *et al.* (2010). Findings showed that both groups obtained similar mean VOT values in the pre-test, which were somewhat closer to their L1 VOT pattern than to the target. In the immediate and delayed post-test, the control group maintained similar mean VOT values. In contrast, the experimental group showed an abrupt increase immediately after the explicit instruction, when they pronounced VOT values longer than the target. This was followed by a decrease in mean VOT values, which brought the values closer to the target, but still longer than those obtained in the pre-test. Therefore, Kupske and Oliveira (2020) concluded that BP speakers benefited from explicit instruction on English VOT, which resulted in more accurate VOT productions.

Contrary to findings from Kupske and Oliveira (2020), Alves and Kampff (2019) did not find positive effects of explicit instruction on the production of English VOT by Brazilian learners of English, although it is important to note that Alves and Kampff (2019) did not use the model by Celce-Murcia *et al.* (2010) but used perceptual training only, which may partially explain such outcome difference. Alves and Kampff (2019) analyzed the long-term effects of perceptual training on the perception and production of English stops by 15 Brazilians divided into experimental and control groups. A pre-test and three post-tests were administered: one Immediate Post-test after the three training sessions that occurred in one week, a Delayed Post-test 1 administered one month after training, and a Delayed Post-test 2 administered three years after training. In order to verify if Brazilians were considering VOT as the main acoustic cue to distinguish English voiceless from voiced stops, the manipulated Artificial Zero VOT was employed in the perception test. Results for perception revealed that the experimental group significantly improved their perception of English stops with Natural and Artificial Zero VOT, while the control group did not. However, accurate perception levels were low for the Artificial Zero VOT, which never surpassed 60% accuracy (perceived as voiced instead of voiceless). Yet, levels for the Natural Zero VOT were high, achieving 95% accuracy. This suggested that

Brazilians had not fully developed competence to use VOT as the main phonetic cue to distinguish sets of English stops. Overall, findings suggested that training had a positive effect on English VOT perception. Regarding the results for production, although the experimental group produced English voiceless stops with longer VOT values in the Delayed Post-test 1 compared to the Pre-test, this increase in VOT values was not maintained in the Delayed Post-test 2. According to the authors, training had a modest contribution to perception, but this contribution was not generalized to production. Based on the findings for perception and production of English VOT by Brazilians, Alves and Kampff (2019) mentioned the need for constant language exposure and use, including high-quality L2 input, as well as the need for consciousness-raising activities with varying degrees of explicitness.

In conclusion, the last sections showed that many studies have analyzed VOT productions by monolingual English and monolingual BP speakers, as well as the perception and production of English VOT by Brazilian learners of English. In general terms, Brazilian learners of English, regardless of the acoustic cue under consideration, show little difficulty in perceiving English voiceless stops as voiceless and English voiced stops as voiced, although VOT, which is the paramount acoustic cue used by native English speakers to do so, seems not to play a major role in identifying and discriminating English stops (Alves; Zimmer, 2015). This minor role played by VOT in Brazilian learners' perception of English VOT seems to be, at least to some extent, extended to production. While monolingual English speakers produce voiceless stops with long VOT values, Brazilian learners of English tend to produce English voiceless stops with VOT shorter than the target, although generally longer than the VOT found in their L1. To help overcome such difficulties, exposure to the target language and instruction are suggested to benefit Brazilian learners in their perception and production of English VOT patterns.

Alves and Kampff (2019) mentioned the lack of exposure to L2 input quantity and quality as a possible explanation for the little or no improvement observed in the perception and production of English VOT patterns by Brazilians. Along with L2 input exposure, it was also mentioned the need for opportunities that allow Brazilian learners to raise consciousness of the target forms. Although their study (Alves; Kampff, 2019) did not show that explicit instruction largely benefits the perception and production of English VOT, results from Kupske and Oliveira (2020) suggest that even a 2-hour explicit instruction session can significantly contribute to English VOT production. Therefore, raising L2 phonological awareness may facilitate the acquisition of English VOT patterns, as Brazilians were shown not to possess high levels of awareness about English VOT (Kivistö-de Souza, 2017).

Following that, the present study aims to obtain a better understanding of L2 phonological awareness by analyzing the possible relation between implicit and explicit L2 phonological awareness and the variables that may affect implicit L2 phonological awareness, namely, explicit L2 phonological awareness, L2 proficiency, and L2 input, given the importance that these variables have shown in affecting linguistic dimensions. Moreover, it is also objected to analyze the levels of implicit and explicit awareness about English VOT in Brazilian learners of English. The next chapter will present the method adopted in the present study.

3 THE PRESENT STUDY

In this chapter, the method to carry out the present study will be addressed in five main sections. The first section refers to the objectives and research questions. The second and third sections are related to the participants and instruments, respectively. The fourth section provides information on the procedures, and the fifth section provides information on the data analysis.

3.1 OBJECTIVES AND RESEARCH QUESTIONS

The present study has three objectives. The first objective is to examine the extent to which Brazilian learners of English possess implicit and explicit knowledge about English VOT. Although some studies have analyzed Brazilians' implicit awareness about English VOT (e.g., Kivistö-de Souza, 2017), to the best of my knowledge, no studies have analyzed their explicit awareness about this English phonetic cue. Therefore, the present study aims to address this gap in the literature and add to the existing knowledge.

The second objective is to examine the relation between implicit and explicit L2 phonological awareness. This objective stems from the need for more empirical studies analyzing the possible relationship between the two types of knowledge, as a means to test the theoretical foundations of the interface positions (Roehr-Brackin, 2024).

The third objective is to examine the variance in implicit L2 phonological awareness about VOT in terms of L2 proficiency, L2 input, and explicit L2 phonological awareness. This stems from the importance of analyzing variables affecting L2 phonological awareness, given its significance for L2 speech development (Roehr-Brackin, 2024). Implicit L2 phonological awareness, and not explicit L2 phonological awareness, was selected to be the outcome variable as implicit knowledge is believed to be the “basic linguistic competence (Ellis, 2004, p. 231), and the core of the discussion on the interface positions seems to be whether implicit knowledge can be impacted by explicit knowledge, the type of knowledge that usually derives from traditional language lessons (Ellis, 2008).

Therefore, three research questions (RQs) and hypotheses (Hs) were elaborated as they are presented below.

RQ 1: To what degree do Brazilian learners of English possess implicit and explicit awareness of English VOT?

H 1: Based on Kivistö-de Souza (2015, 2017), Brazilian learners of English will manifest low degrees of implicit and explicit awareness about English VOT.

RQ 2: Is there a relation between implicit and explicit L2 phonological awareness?

H 2: Given the theoretical foundations of the weak-interface position (Ellis, 2006; Schmidt, 1995) and the empirical findings that implicit and explicit L2 morphosyntactic knowledge are reciprocally associated (Kim; Godfroid, 2023; Suzuki; DeKeyser, 2017), it is hypothesized that there will be a relation between implicit and explicit L2 phonological awareness.

RQ 3: How much of the variance in implicit L2 phonological awareness about VOT can be explained by L2 proficiency, L2 input, and explicit L2 phonological awareness?

H 3: In accordance with the revised Speech Learning Model (Flege; Bohn, 2021) that posits that L2 input affects L2 perception, and based on findings from Kivistö-de Souza (2015), Kim and Godfroid (2023) and Suzuki and DeKeyser (2017), a considerable amount of implicit L2 phonological awareness will be explained by L2 proficiency, L2 input, and explicit L2 phonological awareness.

3.2 PARTICIPANTS

Participants were Brazilian learners of English who spoke Brazilian Portuguese as their first language, recruited online via social media websites (Instagram, Reddit, Facebook, and X). In order to make better generalizations of implicit and explicit English phonological awareness by Brazilians, participants who were currently doing or had already finished an undergraduate or graduate degree in *Letras Inglês* (English Language, Linguistics, and/or Literature studies) were excluded from the data analysis, considering that these Brazilian learners may possess above-average knowledge of English sounds, such as the ones analyzed here.

In total, 42 participants agreed to participate in the study. However, answers from 5 participants were excluded from the data analysis due to the exclusion criteria and inconsistent performance. Out of these 5 participants, three reported studying *Letras Inglês*, one failed to answer more than ninety percent of the instrument measuring implicit phonological awareness, and one other completed the tasks in substantially less time and with inconsistent answers in comparison to the rest of the sample. Therefore, the final sample size consisted of 37

participants, whose demographic and linguistic information is discussed in the following paragraphs and summarized in Table 4.

Table 4 – Participants' background information

Participants' characteristics	Mean (SD)	Range	Participants' characteristics		Frequency (percentage)
Age	27.19 (10.32)	18 - 68	Gender	Male	26 (70.27)
AOL	10.35 (4.03)	3 - 21		Female	11 (29.73)
LexTALE	73.85 (12.23)	51.2 - 96.2	Place of birth (Region of Brazil)	South	23 (62.16)
Time spent abroad (months)	15.73 (47.33)	0.23 - 216		Southeast	10 (27.03)
				Central-West	2 (5.4)
				North	1 (2.7)
Years of English instruction	5.9 (3.84)	0.08 - 15	Occupation	Student	24 (64.86)
				Not-Student	13 (35.13)
Self-estimated general English proficiency (0 = beginner, 10 = advanced)	7.32 (1.33)	3 - 7	Primary English learning type	Language academy	15 (40.54)
				Media	11 (29.73)
				School	8 (21.62)
				Exchange	3 (8.11)
Self-estimated English proficiency: (0 = beginner, 5 = native)			Number of L2s apart from English	Zero	24 (64.86)
Reading	3.73 (0.87)	2 - 5		One L2	7 (18.92)
Listening	3.54 (0.9)	1 - 5		Two L2s	4 (10.81)
Speaking	3 (0.94)	1 - 4		Three L2s	1 (2.7)
Writing	2.95 (0.97)	1 - 5		Four L2s	1 (2.7)

Source: the author.

AOL = Age of Onset of Learning, L2s = second languages, *SD* = standard deviation

Of the 37 participants, 26 were male, and 11 were female. Their mean age was 27.19 years (*SD* = 10.32, range = 18 – 68), and they were born in four out of the five regions of Brazil: 23 in the South, 11 in the Southeast, 2 in the Central-West, and 1 in the North. All of them were currently living in Brazil, except two participants⁹, one of whom, aged 22, was living in Paris,

⁹ In the defense, it was suggested to exclude these two participants who were living abroad, but they were kept here as they were verified not to be outliers.

France, for one year, and the other, aged 35, was living in Glasgow, Scotland, for 18 years. Twenty participants reported having travelled abroad, and their mean length of stay was 15.73 months ($SD = 47.33$, range = 0.23 – 216). Moreover, only 5 participants were left-handed, and none of the 37 participants reported having a hearing problem. Regarding occupation, 24 participants were students of different undergraduate and graduate degree areas, and 13 participants were not studying.

In terms of language-related information, the mean Age of Onset of Learning (AOL) was 10.35 years ($SD = 4.03$, range = 3 – 21). Fifteen participants reported learning English mainly through the language academy, 11 through media (movies, music, TV, games), 8 reported regular school classes, and 3 through exchange programs. Only 3 participants reported never attending English formal classes, and the mean years of English instruction of the 34 participants who reported attending was 5.9 ($SD = 3.84$, range = 0.08 – 15). Moreover, none of the participants had attended a course on English pronunciation.

In relation to the number of additional languages (L2s), 13 participants spoke at least one more L2 apart from English, from these, 7 participants reported speaking one other L2, 4 participants reported two other L2s, one participant reported three more L2s (beginner Spanish, French, and Japanese), and one other participant reported speaking four more L2s apart from English (intermediate Japanese and Italian, and beginner Hungarian and Icelandic). Of these L2s, romance languages were among the most frequent: Spanish ($n = 8$), Italian ($n = 4$), and French ($n = 4$), followed by Korean ($n = 1$), and Hungarian and Icelandic ($n = 1$). The L2 proficiency of the participants ranged from beginner ($n = 12$) to intermediate ($n = 9$).

The self-estimated general English proficiency, on a 10-point Likert scale (0 = beginner, 10 = advanced), of the 37 participants ranged from 3 to 10 ($M = 7.32$, $SD = 1.33$), indicating that self-estimations varied, but concentrated on the higher end of the scale. Regarding specific skills on a 5-point Likert scale (0 = beginner, 5 = native), reading showed the highest mean ($M = 3.73$, $SD = 0.87$, range = 2-5), followed by listening ($M = 3.54$, $SD = 0.90$, range = 1-5), speaking ($M = 3.0$, $SD = 0.94$, range = 1-4), and writing ($M = 2.95$, $SD = 0.97$, range = 1-5).

In the vocabulary measurement called LexTALE used as a proficiency estimation in the present study, participants' mean score was 73.85 ($SD = 12.23$, range = 51.25 – 96.25). Scores from 60 to 80 are considered to resemble upper-intermediate proficiency (Lemhöfer; Broersma, 2012). Of the 37 participants, 5 (13.51%) scored less than 60, 22 (59.46%) scored in this range, and 10 (27.03%) scored above 80. Therefore, participants' proficiency varied from lower-intermediate to advanced.

Participants were asked about the language they use with their five closest people. Most participants reported using only Brazilian Portuguese (BP), although some also use English to varying degrees. No participant reported using only English or another L2 with all the 5 closest persons, and only one participant, currently living in France for one year, reported using another L2 apart from English, which is used with 3 out of the 5 closest persons. Table 5 displays the number of participants and their proportion of BP and English use across the 5 closest persons.

Table 5 – Number of participants and their proportion of Brazilian Portuguese (BP) and English use across their 5 closest persons

Five closest persons	Only BP	More BP	Half BP half English	More English	Only English	Other
Person 1	32	1	2	1	1	-
Person 2	33	2	2	-	-	-
Person 3	33	-	2	-	1	1
Person 4	30	3	1	-	2	1
Person 5	27	2	3	-	4	1

Source: the author.

3.3 INSTRUMENTS

Four instruments were employed in the study. To assess implicit L2 phonological awareness, a perception test called the *Error Detection Test* was adapted from Flege and MacKay (2004) for this study. To assess explicit L2 phonological awareness, a novel questionnaire was designed. To estimate learners' English proficiency, a vocabulary size measure was used. Lastly, to obtain participants' demographic and linguistic information, a background questionnaire was developed.

All four instruments were designed online in the Gorilla Experiment Builder¹⁰ (Anwyl-Irvine *et al.*, 2020) platform. Gorilla Experiment Builder is a user-friendly platform for conducting online research, allowing the creation of complex online tasks and questionnaires. By designing the entire experiment on a single platform, participants can easily carry out the study on their own, without time and place restrictions. This form of study was favored in an attempt to facilitate recruitment, participation, and retention.

The following subsections will approach each instrument closely.

¹⁰ www.gorilla.sc

3.3.1 Implicit L2 phonological awareness

Considering the nature of implicit awareness, unconscious knowledge based on intuition (Ellis, 2004), a perception test called the *Error Detection Test* was designed. The *Error Detection Test* is based on Flege and MacKay's (2004) Experiment 4, in which participants are supposed to identify correct from incorrect samples of a given feature after listening to them. For this study, Brazilian learners of English were presented with English words starting with [p], [t], and [k] (along with distractors), differing in their accuracy of VOT values. The tokens were acoustically manipulated so that their VOT values matched the standard English VOT values found in the literature, whereas the VOT values of the inaccurate tokens were manipulated to resemble the Brazilian Portuguese VOT values found in the literature. In each test trial, participants were first presented with the orthographic word in which the letter representing the target sound was highlighted (in blue and underlined), followed by the audio stimulus. After listening to the target sound, participants had up to 2 seconds to indicate whether the sound was accurately or inaccurately pronounced (Bialystok, 1979). The objective of setting a time limit for each trial was to encourage the use of intuition and thus access their implicit awareness, as there would be very limited time to access explicit knowledge.

3.3.1.1 Stimuli

The stimuli of the *Error Detection Test* consist of the target words and distractors. As for the target words, 15 monosyllabic words were selected, 5 for each voiceless stop ([p], [t], and [k]) in the initial position. Only English words initiating with voiceless stops were selected, as they are generally produced with positive VOT values (Schwartzhaupt; Alves; Fontes, 2015). The aspiration in the English voiceless stops may make the difference between the accurate and inaccurate tokens more noticeable, which lowers the difficulty level of this test. Along with the initial sound, target words were selected considering the following vowels to better control for the phonological context and variance (see Table 6). This way, the target words encompass not only high vowels (the ones reported to increase the VOT values of previous stop consonants (Yavas, 2008; França, 2011)), but also mid and low vowels, in search of generalizing findings to a broader context. Therefore, the high front vowels [ɪ] and [i], along with the high and mid back vowels [u] and [ʊ], and the low front vowel [æ] were chosen to follow the initial voiceless stops, except for the high front vowels following the alveolar [t] that were replaced by mid front vowels [e] to avoid misperception of aspiration by palatalization (Kupske, 2016). Nonetheless,

it was valuable to include high front vowels in the target words as higher vowels seem to increase VOT values and, therefore, make the VOT more prominent in the accurate tokens.

For the distractors, 10 monosyllabic words starting with the [dʒ] sound were chosen. To determine the distractors, words would have to start with a consonant sound represented by a single letter, and this initial sound would have to be suitable for acoustic manipulation to create incorrect-correct samples for each word. Moreover, the initial sound would have to pose some difficulty for Brazilian learners of English. Distractors needed to share these similarities with the target words to facilitate the *Error Detection Test* as participants would have to focus their attention solely on the initial sound of every word presented. Therefore, English words starting with the voiced post-alveolar affricate [dʒ] sound were suitable as they are generally represented by the letters <J> or <G>, which in BP are generally pronounced as a post-alveolar fricative [ʒ] or a palatal fricative [g] in initial position (Seara; Nunes; Lazzarotto-Volcão, 2011).

Table 6 – Target words and distractors included in the *Error Detection Test*

Target words				Distractors	
Initial sound			Following vowel	Initial sound	
[p]	[t]	[k]		[dʒ]	
pig	tell ¹	kid	[ɪ]	job	June
peace	ten ²	keep	[i]	joke	judge
pool	tool	cool	[u]	jump	joint
pack	task	cat	[æ]	jail	gene
pause	tall	call	[ɔ]	joy	gym

Source: the author.

¹ ‘tell’ and ² ‘ten’ are not followed by the high front vowels [ɪ] and [i], which were replaced by mid front vowels [e] to avoid aspiration being perceived as palatalization (Kupske, 2016).

The word frequency of all 25 words was verified at the New General Service List Project¹¹ (NGSL) (Browne; Culligan; Phillips, 2013) in the ‘TOEIC English - NGSL + Sup’ list. The NGSL categorizes words into 7 levels. Level 1 includes the 400 most frequent English words in the corpus (plus 50 supplemental words considered elementary to L2 learners of English), while Level 2 comprises the 401 to 800 most frequent words. The stimuli of the present study are part of Level 1, except for ‘gym’, which is in Level 2.

3.3.1.2 Stimuli recording

¹¹ <https://www.newgeneralservicelist.com/>

Five native speakers of American English recorded the stimuli used in the perception test measuring implicit phonological awareness. To facilitate the recruitment that was online and anonymous, the provision of additional background information, such as the speakers' gender, age, place of birth, and place of current residence, was optional rather than mandatory. Of the 5 speakers, 3 provided additional background information, being two males (one aged 36, born in New York State and currently living in Petersburg, Pennsylvania, and the other aged 64, born and currently living in Florida) and one female (aged 18, born and currently living in Southern California). The two other speakers were identified as one male and one female. The invitation posted on social media made clear that only monolingual speakers of American English were targeted, which was done to avoid the influence of other languages on English VOT values (Kupske; Perozzo, 2024).

Only American English was chosen because this is the English variety that Brazilians are more familiar with. Although it is known that American English consists of a broader term and that the United States is composed of various English accents spoken by different communities, it was not possible to select speakers from a more specific American English variety, such as the standard American English variety from the Midwest.

The speakers recorded the target words embedded in the carrier sentence 'I would say _____', which were read only once. As recruitment was online, speakers recorded themselves with the equipment they had at hand, which caused the audio quality to vary among speakers in some words. Some word productions were considered inappropriate for the test, making it necessary to have four speakers for the test trials and one additional speaker for the familiarization session. For the test trials, one male and one female recorded the target words, and another female and male recorded the distractors. The fifth speaker was a male who recorded the words included in the familiarization session (Appendix A).

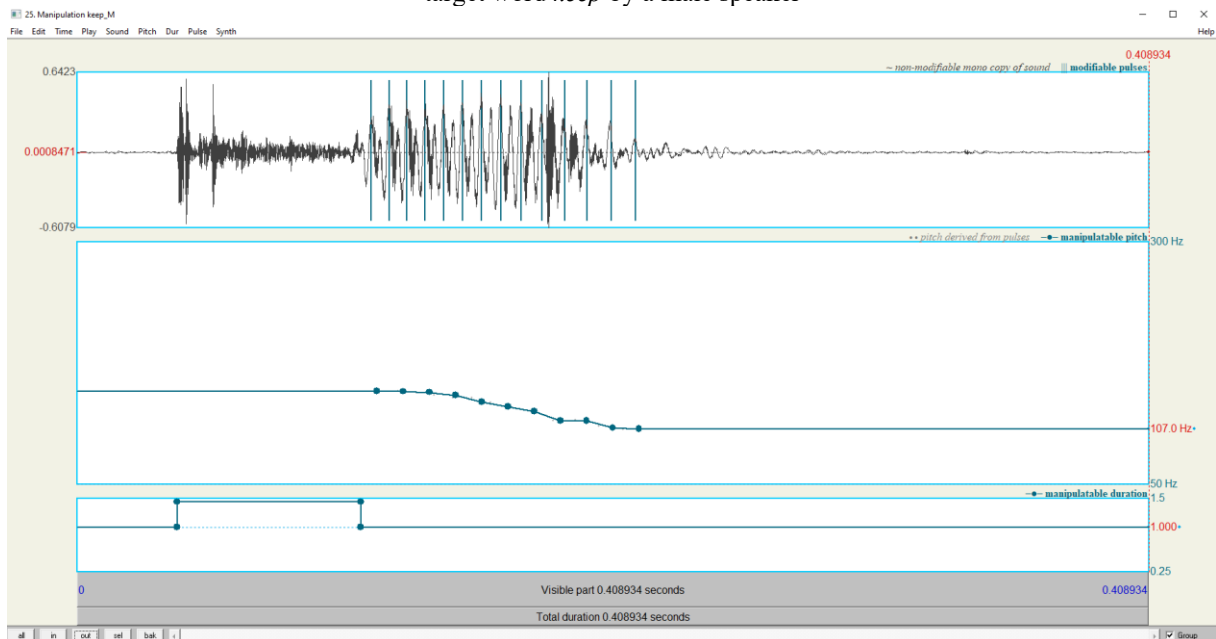
3.3.1.3 Stimuli treatment

The target words and distractors were acoustically manipulated in order to have one accurate and one inaccurate token of each. The target words had their VOT values manipulated so that the accurate token (English-like) displayed the standard American English VOT value of the given stop, as found in the literature, while the inaccurate token (Brazilian-English-like) displayed the BP VOT value found in the literature. For the accurate tokens, the audio recordings by native English speakers were manipulated to match the standard English VOT values as a way to mitigate speaker variability and ensure standard values are displayed in the

perception test, once test-takers are asked about accuracy. VOT manipulation was done using Praat, version 6.3.09 (Boersma; Weenink, 2023), following the *time warping* technique (Schertz; Hawthorne, 2018; Schoonmaker-Gates, 2015; Winn, 2020) to shorten or lengthen the VOT values while preserving the original VOT cues like formant values and pitch. This was done through the duration function inside the Manipulation Editor by adjusting the relative duration within the timing landmarks (see Figure 4). The landmarks were considered the onset of the stop closure and the onset of voicing, adopting Cho and Ladefoged's (1999) procedure.

The mean VOT values used as references for English were obtained from Chodroff and Wilson (2017) in isolated speech, and the values for BP were obtained from Kupske (2016). The standard English VOT values used for the accurate tokens were 89 ms for [p], 98 ms for [t], and 99 ms for [k], whereas the inaccurate tokens followed the standard BP VOT mean values, 15.13 ms for [p], 17.88 ms for [t], and 38.93 ms for [k]. The time difference between VOT values of accurate and inaccurate tokens was 73.7 ms for words starting with [p], 80.12 ms for words starting with [t], and 60.07 ms for words starting with [k].

Figure 4 – The Manipulation Editor window inside Praat showing the duration function (bottom frame) for the target word *keep* by a male speaker



Source: the author.

The VOT duration of this token was 69 ms. Therefore, the relative duration was 1.434 (target duration divided by the actual duration: $99/69 = 1.434$), which is why the line at the bottom frame is leveled up (extending duration) inside the marked dots that are settled at 1 (duration without alteration).

All distractors were also manipulated to have accurate-inaccurate synthetic pairs of all words displayed at the *Error Detection Test*. Similarly to the target words, distractors also had their initial sound manipulated, but instead of voiceless stops, they initiated with the affricate

[dʒ] sound that was replaced by the fricative [ʒ] sound in the inaccurate tokens. To do that, the initial [ʒ] sound of the word ‘genre’ [ˈʒɑnrə] was concatenated with the inaccurate distractors. The accurate tokens, in turn, received the initial [dʒ] sound from the word ‘job’. In Praat (Boersma; Weenink, 2023), the onset and the offset of [dʒ] of each distractor and [ʒ] of ‘genre’ were identified and marked at zero crossing to be separated into two different audio samples each, one sample with the initial sound and the other sample with the rest of the word. The onset of [dʒ] and [ʒ] was regarded as the onset of partial voicing at the beginning of the word production, and the offset as the beginning of the following vowel when there is an abrupt change in intensity or formant frequencies (Clopper *et al.*, 2005). After that, for the creation of inaccurate tokens, the [ʒ] sound was concatenated with an overlap of 10 ms with the remainders of the words using the combine function. The same was done for the accurate tokens, but with the [dʒ] sound from the word ‘job’ (from all words, the [dʒ] of ‘job’ by both male and female speakers was considered a good token of the phoneme). Concatenating the sounds with overlap was preferred over concatenation without overlap to make the synthetic tokens sound more natural, considering the continuum of speech.

3.3.1.4 Test design

A trial of the *Error Detection Test* consisted of two windows: the fixation window, followed by the main window (Figure 5). The fixation window was expected to draw participants’ attention to the stimuli by displaying a centered fixation cross for 1000 ms, followed by a white blank screen for 500 ms. After that, the main window was shown, featuring first the test word with the target sound highlighted in blue and underlined, and the question with the answer options ‘O som sublinhado está correto? F = incorreto / J = correto’ (‘The highlighted sound is correct? F = incorrect / J = correct’)¹². After one second, the audio stimulus is played only once, after which participants have 2000 ms to answer if the indicated sound was correctly pronounced by pressing one of the indicated keys on the computer keyboard. Once the response is given, the next trial is automatically shown. The same happens if no answer is provided within the time limit.

As the reaction time to the responses was obtained, a left-handed version of the test was created with the answer options inverted (F=correct / J=incorrect) since hand dominance may affect reaction time (Dexheimer *et al.*, 2022).

¹² This corresponds to the right-handed version of the test. A left-handed version was also created with the answer options inverted (F=correct / J=incorrect).

Figure 5 – A test trial of the *Error Detection Test*

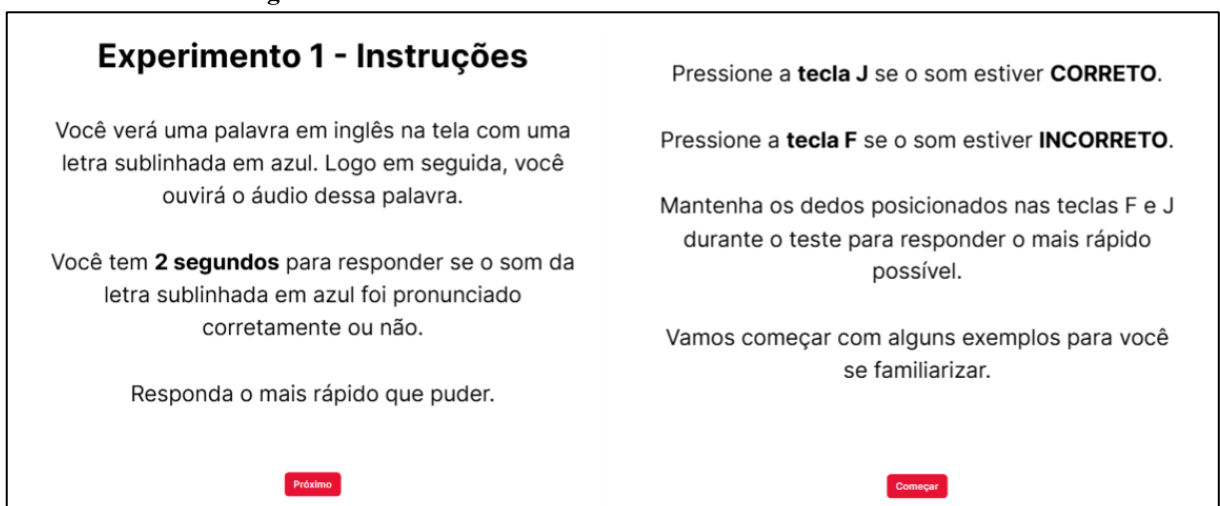


Source: the author.

The sound icon represents the word stimulus and does not appear in the test.

The test consisted of 100 trials, one for each token, with 25 accurate and 25 inaccurate tokens (15 target words and 10 distractors) produced by one female and one male native speaker (Appendix B). Before initiating the test, two windows displaying instructions in PB were shown (Figure 6), followed by a mock test (familiarization session) containing 10 practice trials (Appendix A). Similarly to the test trials, no feedback was given in the familiarization session, and data from the practice trials were not included in the data analysis.

Figure 6 – The two instruction windows of the *Error Detection Test*



Source: the author.

This is the right-handed version of the test. In the left-handed version, the answer options provided on the J and F keys were inverted.

Translation: Experiment 1 – Instructions. You will see a word in English on the screen with one letter underlined in blue. Right after that, you will hear the audio of this word. You have **2 seconds** to decide whether the sound of the underlined letter was pronounced correctly or not. Respond as quickly as possible. Press the **J key** if the sound is **CORRECT**. Press the **F key** if the sound is **INCORRECT**. Keep your fingers on the F and J keys during the test so you can answer as quickly as possible. We will start with a few examples for you to get familiar with.

3.3.2 Explicit L2 phonological awareness

To assess Brazilian learners' explicit L2 awareness, that is, conscious and potentially verbalizable knowledge (Ellis, 2004), a novel questionnaire was designed based on Kivistö-de Souza (2015), Kivistö-de Souza and Lintunen (2023), and Kupisch *et al.* (2014). The questionnaire consisted of two sections with items related to the pronunciation of [p], [t], and [k] in initial position, both in English and BP languages. Distractors were added to prevent participants from recognizing what was being analyzed. The first section of the questionnaire is based on text, while the second section is based on audio. The questionnaire is expected to measure explicit awareness as learners are given unlimited time to consciously think about specific phonetic forms and space to verbalize their answers. Although both this questionnaire and the *Error Detection Test* require learners to focus on specific phonetic forms, they are inherently different. The *Error Detection Test* taps mainly into intuition by giving a limited time for participants to judge the target sound based on 'feeling'. This questionnaire allows participants to relisten to the target sound without a time limit to compare the stimuli to the conceptualization of the sound they have in mind.

Before the questionnaire, a text was presented aimed at bringing participants' attention to individual sounds (Figure 7). This text inclusion was important as participants may be unfamiliar with the concept of individual sounds making up a word, instead of syllables or letters. Moreover, it is important to indirectly clarify to phonetically naïve language users that in the questionnaire, letters are used to represent sounds. To lower the difficulty level of the questionnaire, all items and instructions were in PB.

Figure 7 – Text presented before the questionnaire measuring explicit awareness

Você já parou para pensar quantos sons há em uma palavra?

Tente contar a quantidade de sons que há na palavra em português: 'MESA'. Em 'MESA' temos o mesmo número de sons e de letras, ou seja, quatro. Tente pronunciar cada um separadamente: 'M – E – Z – A'. Porém, nem sempre o número de sons coincide com o número de letras. Veja, por exemplo, a palavra 'HOJE', nela não pronunciamos o 'H' e, então, temos apenas 3 sons ('O – J – E') para 4 letras. Algo parecido acontece com a palavra 'CHÁ', que tem mais letras do que sons. Se você perceber, as letras 'CH' têm apenas um som, seguido pelo som representado pela letra 'Á'. Logo, temos 3 letras e apenas 2 sons: 'CH – Á'.

O mesmo acontece em inglês. Tente contar quantos sons há na palavra 'THREE'. Coincidentemente, na palavra 'THREE' temos 3 sons e não um som para cada uma das 5

letras. Pronunciando em voz alta, você percebe que falamos, primeiramente, o som exclusivo do inglês ‘TH’, seguido do som de ‘R’ e então o som de ‘i’, representado pelas letras ‘EE’. Assim como em ‘CHÁ’, temos mais letras do que sons: ‘TH – R – EE’. Pense então na palavra ‘TAXI’. Nela, temos mais sons do que letras. Prestando bem atenção você percebe que a letra ‘X’ tem 2 sons: ‘K + S’. Então, podemos dizer que ‘TÁXI’ tem 4 letras e 5 sons: ‘T – A – K – S – I’.

Além disso, você já percebeu que cada letra do alfabeto não representa apenas um som, necessariamente?

Se você perceber, a letra ‘S’ na palavra ‘MESA’ tem som de ‘Z’, mas em ‘MÊS’, tem som de ‘S’. Este fenômeno se intensifica se compararmos dois idiomas. Veja: o som da letra ‘A’ da palavra em português ‘MAS’ é bem diferente do som da letra ‘A’ da palavra em inglês ‘BAD’: ‘MAS – BAD’. Outro exemplo é a letra ‘L’. Na palavra em português ‘MAL’, a letra ‘L’ tem basicamente som de ‘U’, enquanto na palavra em inglês ‘MALL’, o ‘L’ tem um som diferente.

Agora que já sabemos a diferença entre sons e letras, vamos responder a um questionário sobre sons do português e do inglês.

Have you ever stopped to think about how many sounds there are in a word?

Try counting the number of sounds in the Portuguese word: ‘MESA’. In ‘MESA’, we have the same number of sounds and letters, that is, four. Try pronouncing each one separately: ‘M – E – Z – A’. However, the number of sounds doesn’t always match the number of letters. Take, for example, the word ‘HOJE’. In this word, we don’t pronounce the ‘H’, so we only have 3 sounds (‘O – J – E’) for 4 letters. Something similar happens with the word ‘CHÁ’, which has more letters than sounds. If you notice, the letters ‘CH’ represent only one sound, followed by the sound represented by the letter ‘Á’. Therefore, we have 3 letters and only 2 sounds: ‘CH – Á’.

The same thing happens in English. Try counting how many sounds there are in the word ‘THREE’. Coincidentally, in ‘THREE’ we have 3 sounds, not one sound for each of the 5 letters. When you say it out loud, you notice that we first pronounce the English sound ‘TH’, followed by the sound of ‘R’, and then the sound of ‘i’, represented by the letters ‘EE’. Just like in ‘CHÁ’, we have more letters than sounds: ‘TH – R – EE’. Now think about the word ‘TAXI’. It has more sounds than letters. If you pay close attention, you notice that the letter ‘X’ has 2 sounds: ‘K + S’. So, we can say that ‘TAXI’ has 4 letters and 5 sounds: ‘T – A – K – S – I’.

What’s more, have you noticed that each letter of the alphabet doesn’t necessarily represent just one sound?

If you look closely, the letter ‘S’ in the word ‘MESA’ has the sound of ‘Z’, but in ‘MÊS’ it has the sound of ‘S’. This phenomenon becomes even clearer when we compare two languages. Look: the sound of the letter ‘A’ in the Portuguese word ‘MAS’ is very different from the sound of the letter ‘A’ in the English word ‘BAD’: ‘MAS – BAD’. Another example is the letter ‘L’. In the Portuguese word ‘MAL’, the letter ‘L’ basically sounds like ‘U’, while in the English word ‘MALL’, the ‘L’ has a different sound.

Now that we know the difference between sounds and letters, let’s answer a short quiz about the sounds of Portuguese and English.

Phonetic transcription of the BP and English words mentioned in the text: MESA /'me.za/, HOJE /'o.ʒi/, CHÁ /ʃa/, MÊS /mes/, MAS /mas/, MAL /maʊ/; THREE /θ.i:/, TEA /ti:/, TAXI /'tæk.si/, BAD /bæd/, MALL /mal/.

3.3.2.1 Text-based section

The questionnaire was divided into two sections: *Text-based* and *Audio-based*. The first section of the questionnaire is called the *Text-based section* and consists of two sets of six statements. The first set of statements compares the initial sound of an English word to the initial sound of a BP word (Figure 8). The second set of statements refers to Brazilians having a characteristic accent when pronouncing the initial sound of English words (Figure 9). In the *Text-based section*, participants are asked to indicate if they agree with the statements. If they agree or not, depending on the set of statements, space for verbalization is given as an attempt to verify the type of knowledge employed. However, not all explicit knowledge is available for verbalization (Ellis, 2004), and yet, answers without a clear mention of VOT can still be considered as somehow deriving from explicit awareness once the statements refer to specific phonetic forms, requiring learners to consciously reflect on this abstract aspect of language for the time needed.

The first set of statements is written as follows, e.g., ‘The pronunciation of the ‘T’ in the English word ‘TO’ (e.g., go to bed) is the same as the pronunciation of the ‘T’ in the Portuguese word ‘TU’ (e.g., Tu és linda)’. Along with the target sound and the word, a short sentence displaying the target word within a context is presented, as the pronunciation of words in isolation may be harder to picture. Among the six statements, three indirectly refer to VOT by comparing pairs of English and BP words that initiate with voiceless stops (one statement for each of the three voiceless stops). As voiceless stops in initial position are generally produced with positive VOT in English and with zero VOT in PB, these three statements are considered wrong. Therefore, higher numbers of correct answers (‘no’) are interpreted as evidence of higher levels of explicit awareness. The other three statements contain pairs of distractors that initiate with either a fricative or an affricate, in which the English word was included in the familiarization session of the *Error Detection Test*. Each word pair was selected regarding the similarity in spelling and/or sound sequence shared between the words, e.g., ‘SALT’ – ‘SAL’ (Table 7). After each statement participants disagreed with (‘no’), a space for verbalization to explain why, if possible, was presented right below.

Figure 8 – Screenshot of the *Text-based section* with the first set of statements

Você concorda com as seguintes frases?

1. A pronúncia do 'T' na palavra em inglês 'TO' (ex.: go **to** bed) é igual à pronúncia do 'T' na palavra em português 'TU' (ex.: **tu** és linda).

☐ Sim

☐ Não

1.1. Explique a diferença entre os sons, se puder:

2. A pronúncia do 'S' na palavra em inglês 'SALT' (ex.: add more **salt**) é igual à pronúncia do 'S' na palavra em português 'SAL' (ex.: adicione **sal**).

☐ Sim

☐ Não

2.1. Explique a diferença entre os sons, se puder:

3. A pronúncia do 'P' na palavra em inglês 'PET' (ex.: my **pet** dog) é igual à pronúncia do 'P' na palavra em português 'PÉ' (ex.: meu **pé** direito).

☐ Sim

☐ Não

3.1. Explique a diferença entre os sons, se puder:

Você concorda com as seguintes frases?

1. A pronúncia do 'F' na palavra em inglês 'FAR' (ex.: go **far** away) é igual à pronúncia do 'F' na palavra em português 'FAZ' (ex.: isso **faz** falta).

☐ Sim

☐ Não

1.1. Explique a diferença entre os sons, se puder:

2. A pronúncia do 'C' na palavra em inglês 'CAR' (ex.: drive a **car**) é igual à pronúncia do 'C' na palavra em português 'CARRO' (ex.: dirijo um **carro**).

☐ Sim

☐ Não

2.1. Explique a diferença entre os sons, se puder:

3. A pronúncia do 'J' na palavra em inglês 'JUNE' (ex.: I arrive in **June**) é igual à pronúncia do 'J' na palavra em português 'JUIZ' (ex.: eu sou **juiz**).

☐ Sim

☐ Não

3.1. Explique a diferença entre os sons, se puder:

Source: The author

The second set of statements is written as follows, e.g., ‘Brazilians have a characteristic accent when pronouncing the ‘T’ in the English words: “**T**ALL”, “**T**OP”, and “**T**ASK”. There is one statement for each voiceless stop and three for distractors, initiating with fricatives, affricates, or nasals (Table 7). As English and PB phonologies embrace different VOT patterns for voiceless stops, it is expected that Brazilians produce voiceless stops with a characteristic accent. Thus, the three statements addressing stops are considered correct. Therefore, higher levels of explicit awareness are seen if participants agree (‘yes’) with the target statements. At the bottom of the pages of the second set of statements, a space for verbalization was given, asking participants to explain, if possible, why they agreed (if so) with the statements above.

Table 7 – Target words and distractors included in the *Text-based section*

First set of statements ¹		Second set of statements	
Target words	Distractors	Target words	Distractors
PET – PÉ	SALT – SAL	PACK – POLL – POP	SAY – SON – SEE
TO – TU	FAR – FAZ	TALL – TOP – TASK	MEET – MOOD – MEN
CAR – CARRO	JUNE – JUIZ	CALL – COLL – CAP	JUST – JURY – JAIL

Source: the author.

¹Word pairs are composed of an English word and a Brazilian Portuguese word.

Figure 9 – Screenshots of the *Text-based section* with the second set of statements

Você concorda com as seguintes frases?

1. Brasileiros têm um sotaque característico ao pronunciar o 'C' nas palavras em inglês: 'CALL', 'COOL' e 'CAP'.

☐ Sim

☐ Não

2. Brasileiros têm um sotaque característico ao pronunciar o 'S' nas palavras em inglês: 'SAY', 'SON' e 'SEE'.

☐ Sim

☐ Não

3. Brasileiros têm um sotaque característico ao pronunciar o 'T' nas palavras em inglês: 'TALL', 'TOP' e 'TASK'.

☐ Sim

☐ Não

Se você concorda com 1., 2. e/ou 3., explique, se puder:

Você concorda com as seguintes frases?

1. Brasileiros têm um sotaque característico ao pronunciar o 'M' nas palavras em inglês: 'MEAT', 'MOOD' e 'MEN'.

☐ Sim

☐ Não

2. Brasileiros têm um sotaque característico ao pronunciar o 'J' nas palavras em inglês: 'JUST', 'JURY' e 'JAIL'.

☐ Sim

☐ Não

3. Brasileiros têm um sotaque característico ao pronunciar o 'P' nas palavras em inglês: 'PACK', 'POOL' e 'POP'.

☐ Sim

☐ Não

Se você concorda com 1., 2. e/ou 3., explique, se puder:

Source: The author

In the *Text-based section*, only monosyllabic words were selected as test words to facilitate the recognition of the target sound. Likewise, aiming to ease the test, the two sets of six statements were separated into four different windows to display only three items at a time, with a space for verbalization right after the statement or at the bottom of the window. The frequency of the English words included in the *Text-based section* was verified at the New General Service List Project (Browne; Culligan; Phillips, 2013). The English words are part of Level 1 in the 'TOEIC English - NGSL + Sup' list, except 'pet', which is in Level 2.

3.3.2.2 Audio-based section

Following the first section of the questionnaire, the *Text-based section*, which is based on statements, the second section of the questionnaire is called the *Audio-based section* and consists of twenty-two items based on audio. Each item is displayed on a single window as follows (Figure 10). On top of the screen, the test word is shown with the first letter highlighted in blue and underlined. Below the word, there is a red button to reproduce the audio stimuli of the word as many times as needed. Below the button, two questions are presented. The first question asks if the highlighted sound in the word is correctly pronounced, and the answer is provided by choosing the 'yes' or 'no' option. No feedback is provided. The second question asks participants to indicate on a 3-point Likert scale (Certain – Semi-certain – Uncertain) how certain they are of their yes/no answer (Kupisch *et al.* 2014). The inclusion of this certainty rating aims to verify whether participants' answers are mostly based on intuition or explicit awareness, by considering that if participants are certain, they are likely basing their answers

on explicit knowledge, whereas if participants are uncertain, they may be relying on intuition (implicit knowledge).

Figure 10 – A screenshot showing one item of the *Audio-based section*

O som sublinhado na palavra em **inglês** foi pronunciado corretamente?

FARM

Play

Reproduza o áudio em 'Play' quantas vezes for necessário.

☐ Sim

☐ Não

Qual é o seu nível de certeza com a sua resposta?

Tenho certeza Um pouco de certeza Não tenho certeza

Próximo

Source: the author.

The red box indicating that the audio can be played as many times as needed was displayed only on the first item.

Twenty-two words were included in the *Audio-based section* (Table 8), one in each item. Half the words were in English and the other half in BP. Among the eleven words in each language, nine were target words starting with voiceless stops, and the other three were distractors. Only monosyllabic words and English words that are part of Level 1 in the ‘TOEIC English - NGSL + Sup’ list were selected (Browne; Culligan; Phillips, 2013). Among the nine target words in each language, six (two for each voiceless stop) displayed inaccurate VOT values, and three (one for each voiceless stop) displayed accurate VOT values.

Table 8 – Words included in the *Audio-based section*

	English					Brazilian Portuguese				
	Target words			Distractors		Target words			Distractors	
Initial sound	[p]	[t]	[k]	[f]	[ʒ]	[p]	[t]	[k]	[z]	[ʒ]
Accurate	pen	take	cold	farm	-	pai	tá	cai	-	já
Inaccurate	poor	talk	can	-	joy	pé	tem	que	sim	-
	put	time	cut			pós	tom	cão		

Source: the author.

Regarding stimulus, one female native speaker of American English and one female native speaker of BP, both monolingual, recorded the stimuli for the *Audio-based section*. The American speaker did not provide additional background information, such as age, place of birth, or place of current residence. The BP speaker was 24 years old, born and currently living in Florianópolis, Santa Catarina. The tokens were embedded in the carrier sentence ‘I would say: _____’/‘Eu diria: _____’ and read only once.

In relation to stimuli treatment, the VOT of the accurate and inaccurate tokens was acoustically manipulated to present accurate and inaccurate VOT values, respectively, similarly to what was done for the *Error Detection Test*. The accurate tokens displayed the standard VOT values of the given language according to the literature. The English words displayed values from Chodroff and Wilson (2017) in isolated speech, whereas the BP words displayed values from Kupske (2016). The opposite was done in the inaccurate tokens. The inaccurate English tokens displayed the standard BP VOT values from Kupske (2016), and the inaccurate PB tokens display the VOT values from Chodroff and Wilson (2017) in isolated speech. Acoustic manipulation was done in Praat, version 6.3.09 (Boersma; Weenink, 2023), using the same procedures described in *Section 3.3.1.3*.

3.3.3 Proficiency test

Considering that L2 phonological awareness may be substantially affected by L2 proficiency, the present study employed a test to estimate Brazilian participants’ general English proficiency. The test selected to be included here is the Lexical Test for Advanced Learners of English, called LexTALE¹³ (Lemhöfer; Broersma, 2012). LexTALE is a short yes/no English vocabulary test developed specifically for psycholinguistics studies (Lemhöfer; Broersma, 2012). In the test, participants are asked to indicate if each of the 60 items (40 words, 20 nonwords) presented one by one in a fixed order is an existing English word or not. The target population is advanced English learners who started learning English around the age of 10 in instructed language learning settings. Yet, the test can be used to discriminate between lower intermediate, upper intermediate, and advanced proficiency levels. The test is unspeeded and takes an average of 3.5 minutes to complete (Lemhöfer; Broersma, 2012). Therefore, this vocabulary test is considered a rapid and easy measure of general English proficiency.

¹³ www.lextale.com

In a study correlating LexTALE scores with other measures of L2 proficiency, such as self-rating and word translation performance, LexTALE was validated as a good predictor of English vocabulary knowledge (Lemhöfer; Broersma, 2012). The large-scale study tested Dutch and Korean speakers of English as an L2 and found substantial correlations between the two speaker groups with distant L1 backgrounds, which suggested the suitability of the test for a variety of language backgrounds. LexTALE scores from both groups were found to correlate substantially with a measure of general English proficiency and to be generally better than self-ratings. Therefore, LexTALE seemed a valid test to assess general English proficiency rapidly and easily.

This lexical test was selected because, different from other tests, LexTALE is validated as an L2 vocabulary measurement that provides a fair indication of general English proficiency (Lemhöfer; Broersma, 2012). Moreover, the test is free and adjustable to the online platform used to develop the other instruments of the present study. Therefore, the data collection can occur on a single page, facilitating the study. Lastly, this test was suitable for the proficiency targeted here, as participants were expected to be intermediate to advanced English learners.

LexTALE scores are the percentage of correct answers, corrected for the unequal proportion of words and nonwords (Lemhöfer; Broersma, 2012). Therefore, it ranges from 0 to 100. The mean score obtained in the study examining Dutch and Korean advanced English learners was 70.7 (Lemhöfer; Broersma, 2012). According to their results, scores from 60% to 80% resemble upper-intermediate proficiency.

3.3.4 Background Questionnaire

To assess participants' linguistic and demographic variables and obtain measures for input, participants were asked to answer a questionnaire addressing personal and language-related background information based on Kim (2020), Kupske (2016), and Mendes (2017). The background questionnaire (Appendix C) was divided into three parts. The first part of the questionnaire consisted of personal background information, such as age and occupation, and language learning information, such as AOL. The second part of the questionnaire assessed participants' L2 input with items addressing different contexts in which they might use English. The third part of the questionnaire involved additional questions regarding English proficiency, additional languages, and pronunciation instruction.

The first part of the questionnaire contained thirteen main questions, and apart from typical personal questions, Question 8 (Q8) asked participants if they had any hearing problems.

This question was included to ensure that only data from learners without hearing difficulties was included in the analysis, as the *Error Detection Test* contained audio stimuli. Q9, about hand dominance, was included to direct participants to the right- or left-handed version of the *Error Detection Test*.

The second part of the questionnaire contained one main question (Q14) asking participants to indicate the number of hours they spend on a typical week doing activities in English. Such activities encompass English input in the four language skills (speaking, listening, writing, and reading) and are distributed in 18 items. For each item, there is a 6-point Likert scale to indicate the number of hours spent on the activity mentioned. This type of question was favored in place of those general indications of language use, across language skills and places, for example, as participants may not remember the contexts/activities in which they use English. Moreover, general indications may be somewhat abstract to language users who are not familiar with linguistic studies. Therefore, asking for the number of hours instead of using general adverbs of frequency, like ‘always’ and ‘never’, may result in a more reliable comparison of the amount of L2 input among participants.

The third part of the questionnaire aimed to obtain additional language information. Q15 asked participants to think about the five closest people they are in contact with and to indicate the language they use to communicate. This question was based on Q60 of Kupske’s (2016) sociolinguistic questionnaire. From Q16 to Q18, participants were asked about their English proficiency levels in different skills. Q19 asked participants if they speak other languages, and Q20 asked if they had taken any courses on English pronunciation, as explicit instruction may affect explicit knowledge (Ellis, 2008). The last question (Q21) was open and asked participants if they had additional comments regarding their English use that might be relevant to the study.

3.4 PROCEDURES

Data collection occurred online in five main steps. Participants were recruited online via social media websites (Instagram, Reddit, Facebook, and X). The study link was made available with the invitation post, so Brazilian learners of English interested in participating could access the study at any time. The invitation post was shared several times by different people, and anyone willing to voluntarily participate in the study was able to take the experiment at their discretion. Therefore, the full scope of the recruitment is unknown.

Once participants clicked on the study link, the consent form with an overview of the study was presented. The consent form made clear that participation was voluntary, along with

the other requirements of the *Comitê de Ética em Pesquisa com Seres Humanos* (CEPSH/UFSC)¹⁴.

If participants agreed to voluntarily participate, they were automatically led to the first instrument, the Background questionnaire. The Background questionnaire was the first instrument because Q9 directed participants to the version of the *Error Detection Test* of their hand dominance. Participants took an average of 06 minutes to read the consent form and 08 minutes to complete the Background questionnaire.

The second instrument was the *Error Detection Test*, which measured implicit phonological awareness, followed by the questionnaire that measured explicit awareness. The *Error Detection Test* was presented first to prevent participants from recognizing what was being analyzed before this task, which could occur during the questionnaire. Consequently, participants would likely heavily depend on their conscious knowledge about VOT to complete this test, which is intended to tap into intuition. The second and third instruments lasted an average of 09 minutes and 25 minutes, respectively.

The fourth and last instrument was the English vocabulary test used to estimate English proficiency. This test was left to the end in case participants got tired during the data collection session, so that tiredness and lack of attention would not affect the results of the main instruments measuring awareness. Also, by arranging the instruments in this way, there was an alternation between the types of instruments, a questionnaire followed by a test, making the study more dynamic and less repetitive. The LexTALE lasted around 04:30 minutes.

In total, participants took an average of 48 minutes to complete the data collection session.

3.5 DATA ANALYSIS

Scores for each variable of the study were calculated according to their corresponding instrument. The next four subsections approach the analysis of each instrument separately. The fifth and last subsection addresses the statistical procedures involved in the analysis to answer the three research questions.

3.5.1 Implicit awareness score

¹⁴ The project was approved on June 20th, 2025, under the CAAE (*Certificado de Apresentação para Apreciação Ética*) 88371925.9.0000.0121.

For implicit L2 phonological awareness, an *Implicit Awareness Score* ranging from 0 to 100 was calculated based on the percentage of correct responses from the *Error Detection Test*. The data from the Gorilla Experiment Builder¹⁵ (Anwyl-Irvine *et al.*, 2020) provides information on several metrics such as response, accuracy, and reaction time¹⁶. This score is the percentage of correct responses to the 60 test words (30 accurate and 30 inaccurate samples), disregarding the 40 distractors. Therefore, the number of correct responses to the 60 test words was divided by 60 and multiplied by 100 to calculate the percentage of correct responses per participant. This way, the tokens that were not indicated as correct or incorrect within the time limit (timeouts) were considered incorrect responses in the calculation, and the higher the score, the higher the implicit L2 phonological awareness level. Apart from a single *Implicit Awareness Score* encompassing VOT of the three English voiceless stops [p t k], scores were also calculated for each voiceless stop separately, following the same calculation, but instead of dividing by 60, the score for each stop was divided by 20, which is the number of tokens for stop (10 accurate and 10 inaccurate samples). This was done to verify the possible difference in awareness among the three stops.

Moreover, two other calculations were done using the data obtained from the *Error Detection Test*. To check if rejecting inaccurate tokens was more difficult than accepting accurate ones, as observed by Kivistö-de Souza (2017), the percentage of correct responses per token accuracy (accurate and inaccurate) was calculated by summing the correct responses to the 30 accurate and 30 inaccurate (10 per stop) test tokens, dividing by 30, and multiplying the outcome value by 100. Lastly, the number of timeouts per participant was also calculated to remove possible participants who failed to answer more than 90% of the task. The total number of timeouts to the 60 test words for the 37 participants (2.220 tokens in total) was 52, divided among 18 participants ($M = 0.92$, $SD = 1.21$, range = 0 – 5). The one participant who got 7 timeouts in the test words (more than 10%) was excluded from the analyses and is not included in any results reported here.

3.5.2 Explicit awareness score

¹⁵ The *Error Detection Test*, the *Audio-Based section* of the questionnaire measuring explicit awareness, and the LexTALE were designed using the ‘Task Builder 2’ function of the platform, whereas the *Text-Based section*, the Background questionnaire, and the consent form were designed using the ‘Questionnaire Builder 2’.

¹⁶ Due to the limited time, the reaction time obtained in the *Error Detection Test* could not be analyzed in the present study and will be reported in future publications.

For explicit L2 phonological awareness, a composite score called the *Explicit Awareness Score* was calculated based on answers from the two parts of the questionnaire targeting explicit awareness. As this is a novel questionnaire, the internal consistency of all yes-no items was verified to find a way of gathering the different types of items (Appendix D). The questionnaire contains two sections, each of which has two different sets of statements (*Text-based section*) or items (*Audio-based section*). The *Audio-based section* contains eighteen items related to VOT, of which nine contain English words, and nine contain BP words (3 displaying accurate VOT values and 6 displaying inaccurate VOT values). The average inter-item correlation of the four types of items, accurate and inaccurate English and BP items, was calculated separately (four calculations), and the only type of item that demonstrated consistent measurement of the same construct was the 6 items in English displaying inaccurate (BP-like) VOT values. For these 6 items together, the average inter-item correlation was .15 (95% CI [.07, .23]). Similar calculations were done for the *Text-based section*.

The *Text-based section* contains six statements referring to VOT, three comparing the initial sound of an English word to a BP word, and three alluding to Brazilians having a characteristic accent when pronouncing English initial sounds. Therefore, both sets of statements were tested for internal consistency to verify if they measure the same construct, namely, explicit L2 phonological awareness. The average inter-item correlation of the first set of statements was .06 (95% CI [-.15, .30])¹⁷, whereas the average inter-item correlation of the second set of statements was .21 (95% CI [.01, .42]), which shows that the second but not the first set of statements consistently measured the same underlying construct. This may be partially explained by the high variance of correct responses found among the statements of the first set, in which the statement for [t] received 26 correct responses ('no'), while the statements for [k] and [p] received only 6 and 2 correct responses, respectively. In the second set, the same stops, [t], [k], and [p], received 20, 13, and 11 correct responses ('yes') each, respectively. This seems controversial, as the first set of statements, which compares specific sounds between languages, is apparently less abstract than the second set of statements, which relates to the Brazilian accent in the initial sound of different English words. Moreover, although most participants agreed that the [k] and the [p] sounds are the same in English and in BP, some also agreed that Brazilians have a characteristic accent when pronouncing these English sounds. This controversy leads to the presumption that some participants were answering the second set of statements considering other phonetic aspects of the words apart from VOT, such as the

¹⁷ The statement related to the [t] sound in the first set correlated negatively with the two other statements of the set.

Brazilian accent when pronouncing the following sounds after the voiceless stop. This presumption was partially confirmed for some participants who mentioned other aspects of the words instead of the initial sound in the open-ended responses provided, e.g. “Eu acho que, quando não tomamos cuidado "mood" vira "moodi" e "meat" vira "miti"”¹⁸ (here, the vowel epenthesis in word ending is mentioned), and “Acho que o sotaque brasileiro nessas palavras não está ligado ao som da letra inicial e sim à pronúncia do resto da frase”¹⁹. Nevertheless, some participants did indirectly refer to VOT in their open-ended responses, e.g., “A pronúncia do "C" é mais seca quando alguém (iniciante brasileiro) está aprendendo inglês.”²⁰, and “Quanto a 1. e 3., os brasileiros tendem a não aspirar as letras.”²¹.

The *Explicit awareness score* was computed using the 6 audio items in English displaying inaccurate VOT values from the *Audio-based section*, plus the 3 statements of the first and the second set of statements referring to VOT from the *Text-based section*, comprising 12 items in total. The first set of statements was included in this score despite showing a low average inter-item correlation value (.06), as these three statements are considered to be more directly related to VOT than the statements of the second set, which reached a moderate average inter-item correlation value (.21). Like the second set of statements, the 6 audio items in English with inaccurate VOT also showed moderate internal consistency, which is the reason they were both included in this score. The *Explicit Awareness Score* is the percentage of correct responses to the 12 items, ranging from 0 to 100. The higher the score, the higher the explicit L2 phonological awareness level.

Apart from computing the *Explicit Awareness Score*, the percentages of correct responses for each item type²² from both sections of the questionnaire were also obtained. Additionally, the frequency and relative frequency (in percentage) of the certainty ratings from the *Audio-based section* were calculated for each item type separately (Appendix E). For each item type, the frequency and relative frequency of the three certainty ratings (Certain – Semi-certain – Uncertain) were calculated for the correct and incorrect answers, as well as for all items of the group together (correct + incorrect answers). This calculation enables the analysis of how certain participants were of their answers to each item type in relation to the accuracy

¹⁸ “I think that when we're not careful, "mood" turns into "moodi" and "meat" turns into "miti".”

¹⁹ “I think the Brazilian accent in these words is not related to the sound of the initial letter, but rather to the pronunciation of the rest of the sentence.”

²⁰ “The pronunciation of the letter "C" is drier when someone (a Brazilian beginner) is learning English.”

²¹ “Regarding 1. and 3., Brazilians tend not to aspirate the letters.”

²² The four item types in the *Audio-based section* are the audio items in BP with accurate VOT and with inaccurate VOT, and the audio items in English with accurate VOT and with inaccurate VOT. The two item types in the *Text-based section* are the first set of statements and the second set of statements.

of their responses, i.e., how certain they were of their correct and incorrect answers to each item type. For the calculation of the relative frequency in percentage, the frequency of each rating (for each item type and answer accuracy) was divided by the number of items (111 accurate items in BP and in English, and 222 inaccurate items in BP and in English) and multiplied by one hundred.

To verify the possible difference of responses within the three voiceless stops separately, the frequency and relative frequency of correct responses to each statement of the *Text-based section* were obtained. For the *Audio-based section*, the mean percentage of correct responses to the two audio items in English with inaccurate VOT for each stop, the ones included in the *Explicit Awareness Score*, was obtained by scoring one point for each correct response, summing the two items of each stop, and dividing by two. The outcome value was then multiplied by 100 to obtain the percentages that were either 0%, 50%, or 100%.

Finally, the open-ended questions from the *Text-based section* were analyzed to verify what participants were considering when answering the statements related to VOT.

3.5.3 L2 input score

To examine whether English input explains part of the variance in implicit phonological awareness of English VOT, a composite score for English input, referred to as the ***L2 Input Score***, was calculated using 13 of the 18 items from Q14 that demonstrated reliable internal consistency. Q14 from the Background questionnaire consisted of 18 items related to the amount of time spent (on a 6-point Likert-scale) on activities in English, of which 11 items were about oral skills (listening and speaking) and 7 about written skills (reading and writing). The internal consistency of the items, inputted in different combinations, was tested using Cronbach's Alpha (Appendix F). The Cronbach's Alpha value of the 11 items related to oral skills was .49 (95% *CI* [.22, .68]), with four items that correlated negatively with the scale (Q14.6, Q14.7, Q14.8, Q14.11). After removing these four items, the new Cronbach's Alpha value for the remaining 7 items related to oral skills was .61 (95% *CI* [.37, .78]). The Cronbach's Alpha value for the 7 items related to written skills was .65 (95% *CI* [.47, .77]), with item Q14.13 correlating negatively with the scale. With this item removed, the new Cronbach's Alpha value for written skills was .70 (95% *CI* [.55, .81]). Finally, the Cronbach's Alpha value for all 18 items together was .71 (95% *CI* [.57, .82]), with the same four items of the oral skill set correlating negatively with the scale (Q14.6, Q14.7, Q14.8, Q14.11). After removing these

four items, the new value for the remaining 14 items was .76 (95% *CI* [.63, .85]). Of these 14 items, when item Q14.14 was removed, the new Cronbach's Alpha value was elevated to .77 [.64, .86]. Therefore, the remaining 13 items from Q14 were used to compute the *L2 Input Score* due to their reliable internal consistency. The responses to the 6-point Likert scales were converted to values from 1 to 6 and summed to create a composite score, ranging from 13 to 78. The higher the score, the higher the amount of English input.

3.5.4 L2 proficiency score

To examine whether English proficiency would explain part of the variance in implicit phonological awareness about English VOT, the *L2 Proficiency Score* for English proficiency was obtained from the LexTALE test. The LexTALE score consists of an averaged percentage of correct answers, corrected for the unequal proportion of words and nonwords (Lemhöfer; Broersma, 2012). Therefore, the score ranges from 0 to 100. The calculation is as follows: $((\text{number of words correct} / 40 * 100) + (\text{number of nonwords correct} / 20 * 100)) / 2$. The higher the score, the higher the English proficiency.

3.5.5 Research questions

The three research questions (RQs) were answered using descriptive and inferential statistics, conducted in the JASP statistics program, version 0.19 (JASP Team, 2025). RQ1, which asks the degree to which Brazilian learners of English possess implicit and explicit awareness of English VOT, was answered with the descriptive statistics of the *Implicit Awareness Score* and the *Explicit Awareness Score*. For RQ2, regarding the relation between implicit and explicit L2 phonological awareness, a correlation analysis was run. As for the testing of correlation assumptions, the *Implicit Awareness Score* did not fit a normal distribution (*p-value* of Shapiro-Wilk <.001). Therefore, the non-parametric Spearman's rank correlation was used. The statistical power of the correlation was calculated in RStudio (Posit team, 2025), version 4.5.2, using the 'pwr' package (Larson-Hall, 2015, p. 106). RQ3, which refers to how much of the variance in implicit L2 phonological awareness about VOT can be explained by L2 proficiency, L2 use, and explicit L2 phonological awareness, was answered through a standard multiple regression analysis.

This section described the method of the study. The next section will present the results and discussion

4 RESULTS AND DISCUSSION

This chapter reports the results for each research question and presents a discussion of the findings. The first subsection addresses the first RQ, which asks about Brazilian learners' levels of implicit and explicit awareness of English VOT. The second subsection seeks to answer the second RQ by addressing the possible relationship between implicit and explicit L2 knowledge. The third subsection addresses the third RQ, about factors affecting implicit L2 knowledge. The discussion of the findings is presented in each section, after the results are reported. Finally, the fourth section presents a general discussion of the findings.

4.1 RQ1: BRAZILIAN LEARNERS' LEVELS OF IMPLICIT AND EXPLICIT AWARENESS OF ENGLISH VOT

Research Question 1 investigates the degree to which Brazilian learners of English exhibit implicit and explicit awareness of English VOT. The hypothesis, based on Kivistö-de Souza (2015, 2017), was that Brazilian learners would possess low levels of both implicit and explicit awareness of English VOT. To answer RQ1, we investigated the descriptive statistics of the *Implicit Awareness Score* and the *Explicit Awareness Score*, as well as the complementary data obtained from the *Error Detection Test* and the questionnaire measuring explicit awareness. Plots and tables were created to assist in the analysis of data and are distributed throughout the section. Appendix G provides the individual *Implicit Awareness Score* and *Explicit Awareness Score* per participant.

Starting with Brazilian learners' implicit awareness of English VOT, the mean *Implicit Awareness Score* was 53.15 ($SD = 5.7$, range = 43.33 – 70), which is slightly above the chance level (50%). The standard deviation and the range suggest that participants differ in their implicit awareness of English VOT. This is confirmed in the boxplot of Figure 13. The orange box for the *Implicit Awareness Score* has an interquartile range (IQR) of 5, which is placed between the 50% and 60% gridlines. Scores are concentrated slightly above the change level, with two outliers near the 70% gridline.

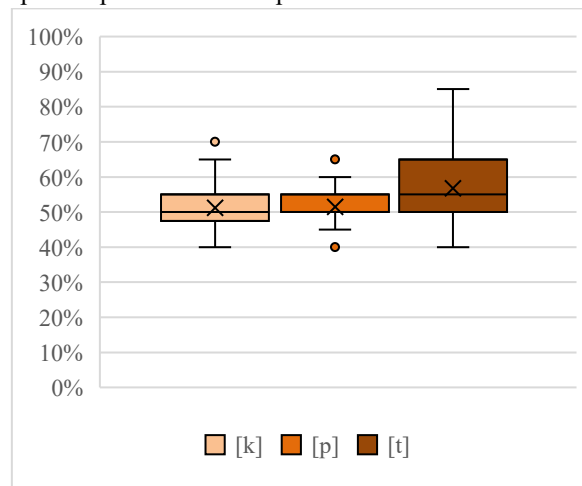
Table 9 – Descriptive statistics of the *Implicit Awareness Score* and the *Explicit Awareness Score*

Score	N	Mean	SD	Range
<i>Implicit Awareness Score</i>	37	53.15	5.7	43.33 – 70
<i>Explicit Awareness Score</i>	37	30.63	15.22	0 – 58.33

Source: the author.

In relation to the implicit awareness of the English VOT for each English voiceless stop separately, participants seemed to be consistent in their answers across the three stops in the *Error Detection Test*. The mean percentage of correct responses to the 20 items initiated with a velar [k] was 51.22% ($SD = 6.39$, range = 40 – 70), whereas for the bilabial [p] and the alveolar [t] it was 51.49% ($SD = 5.12$, range = 40 – 65) and 56.76% ($SD = 10.36$, range = 40 – 85), respectively. Figure 11 displays the boxplot for the percentage of correct responses per voiceless stop, where the box for the alveolar [t] is slightly longer, surpassing the gridline at 60%, than the boxes for the velar [k] and bilabial [p], which are below the gridline at 60%. This graphically demonstrates the higher mean value and standard deviation found for the alveolar [t] in comparison to the velar [k] and bilabial [p].

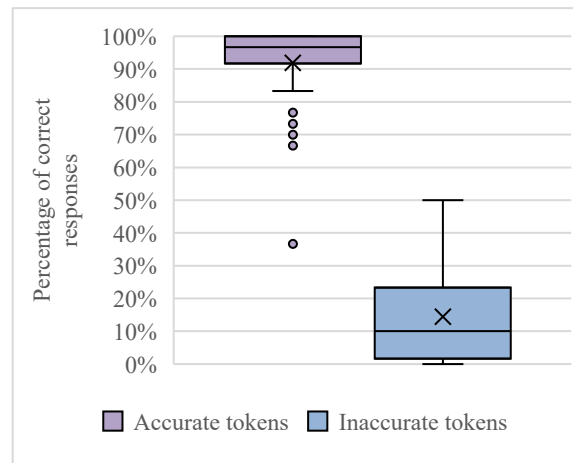
Figure 11 – Boxplot showing the percentage of correct responses per voiceless stop from the *Error Detection Test*



Source: the author.

Still talking about the results from the *Error Detection Test*, rejecting inaccurate tokens was substantially more difficult than accepting accurate ones, following the observations in Kivistö-de Souza (2017). The mean percentage of correct responses to the accurate tokens was 91.89% ($SD = 12.75$, range = 36.67 – 100), whereas to the inaccurate tokens it was 14.41% ($SD = 13.72$, range = 0 – 50). The range for the correct responses to the inaccurate tokens (0 – 50) shows that even the best participants were at chance level, as no one scored above 50% – indicating more than 50% of the inaccurate tokens were incorrect. The boxplot of Figure 12 displays the boxes for the percentage of correct responses to the accurate and inaccurate tokens, in which the former is located at the top of the plot, above the gridline at 90%, whereas the latter is located near the bottom of the plot, below the gridline at 30%.

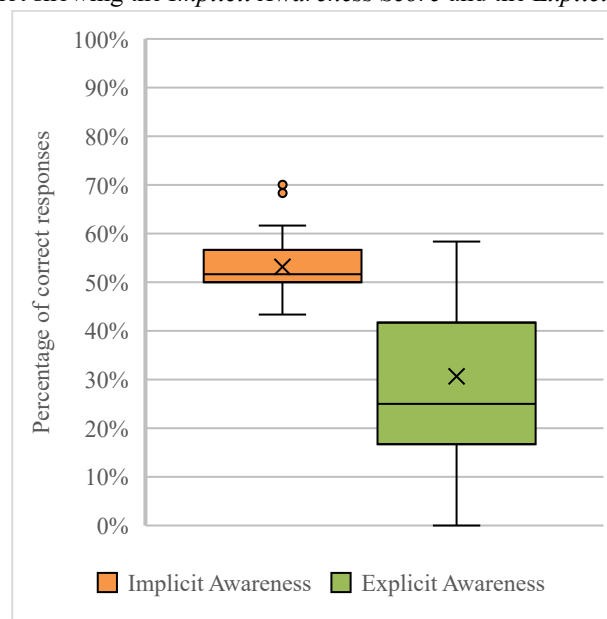
Figure 12 – Boxplot showing the percentage of correct responses to accurate and inaccurate tokens from the *Error Detection Test*



Source: the author.

Regarding Brazilian learners' explicit awareness of English VOT, their mean *Explicit Awareness Score* was 30.63 ($SD = 15.22$, range = 0 – 58.33). The considerably high standard deviation, which is half the mean score, represents the high variance of answers to the questionnaire items among participants. As visually represented in the boxplot of Figure 13, the IQR of the green box for the *Explicit Awareness Score* is 25. The lower quartile is slightly below the 20% gridline, while the upper quartile is slightly above the 40% gridline, with whiskers going from 0% to almost 60%. When compared to the orange box for the *Implicit Awareness Score*, the box for the *Explicit Awareness Score* is located below on the scale, with a considerably higher IQR.

Figure 13 – Boxplot showing the *Implicit Awareness Score* and the *Explicit Awareness Score*



Source: the author.

As for the item types from the *Audio-based* and *Text-based* sections of the questionnaire, the percentage of correct responses varied considerably. For the item types from the *Audio-based section*, as displayed in Table 10, accurate items were easily identified as correct, while inaccurate items were considerably harder to identify as incorrect, in both Brazilian Portuguese and English. The minimum value of the correct responses to the accurate items in both languages resembles the maximum value of the correct responses to the inaccurate items in BP and English. The standard deviations for the item types with inaccurate VOT are considerably higher than those for the item types with accurate VOT, which indicates that participants were more consistent in their answers when listening to accurate items, although part of this difference may be accounted for by the difference in the number of items of each type (3 accurate and 6 inaccurate items).

Table 10 – Percentage of correct responses to the accurate and inaccurate items from the *Audio-based section*

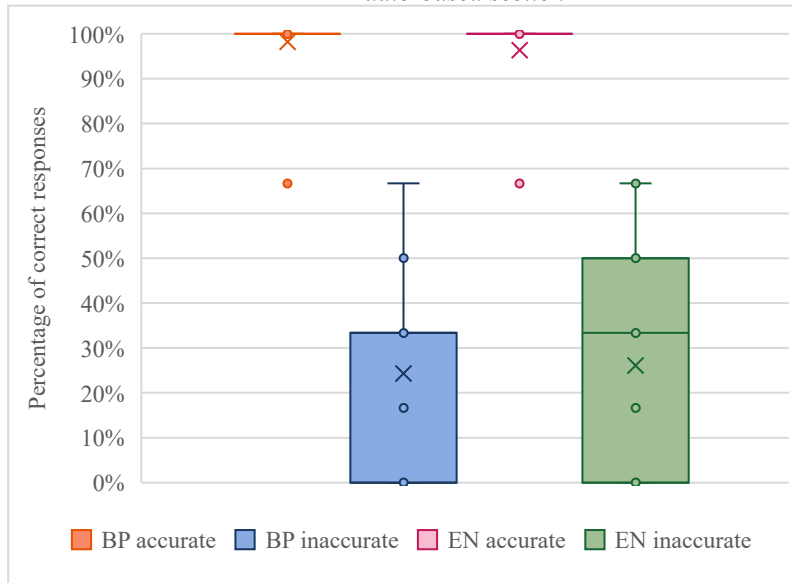
	Accurate items			Inaccurate items		
	Mean (<i>SD</i>)	Range	<i>N</i>	Mean (<i>SD</i>)	Range	<i>N</i>
BP	98.2 (7.64)	66.67 – 100	111	24.32 (19.48)	0 – 66.67	222
English	96.4 (10.49)	66.67 – 100	111	26.13 (22.06)	0 – 66.67	222

Source: the author.

With regards to the *Text-based section*, the mean percentage of correct responses to the first set of statements was 30.63% ($SD = 21.34$, range = 0 – 100), whereas to the second set of statements it was 39.64% ($SD = 33.18$, range = 0 – 100). Again, the high SD s indicate high variability within the answers.

As additional information, the boxplots of Figures 14 and 15 visually represent the IQRs of the six item types. In Figure 14, the IQRs of the audio items with accurate VOT are equal to 0 and located at the upper end of the scale, while the IQRs of the audio items with inaccurate VOT are equal to 33.34 for BP and 50 for English, located at the bottom of the scale. In Figure 15, the IQRs of the first and second sets of statements are 0 and 66.67, respectively. For the first set of statements, the percentage of correct responses is concentrated at 33.34%, meaning most participants got one correct answer out of three statements; the other percentages of correct responses (0%, 66.67%, and 100%) are taken as outliers given their low frequencies.

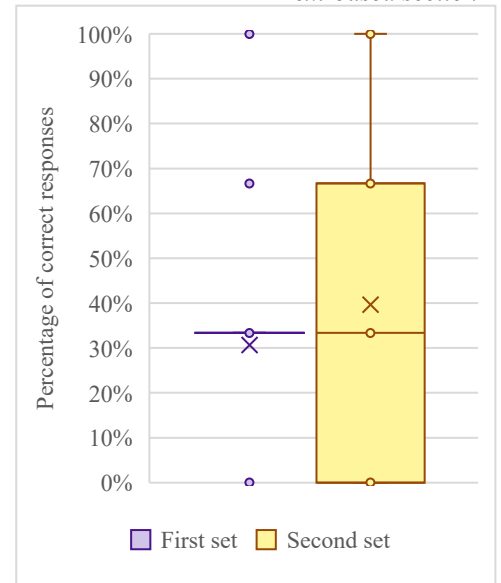
Figure 14 – Boxplot for the percentage of correct responses per item group (BP accurate and inaccurate, and English accurate and inaccurate) from the *Audio-based section*



Source: the author.

Key: BP = Brazilian Portuguese; EN = English.

Figure 15 – Boxplot for the percentage of correct responses to the first and second sets of statements from the *Text-based section*



Source: the author.

In accordance with the *Error Detection Test* (Figure 11), in the questionnaire measuring explicit awareness, participants performed generally better in items for the alveolar [t] compared to the velar [k] and bilabial [p]. In the first set of statements of the *Text-based section*, 26 participants (70.27%) correctly answered the statement related to [t], whereas only 6 (16.22%) and 2 (5.40%) participants correctly answered the statements related to [k] and [p], respectively. In the second set of statements, the number and percentage of correct answers to the statements related to [t], [k], and [p] were 20 (54.05%), 13 (35.13%), and 11 (29.73%), respectively. For the *Audio-based section*, the audio items in English with inaccurate VOT, the ones included in the *Explicit Awareness Scores* given their reliable internal consistency (average inter-item correlation = .15 (95% CI [.07, .23]), also displayed higher mean percentages of correct answers to the items related to the alveolar [t], 47.3% ($SD = 40.73$), in comparison with the items related to the velar [k], 5.4% ($SD = 15.74$), and bilabial [p], 25.68% ($SD = 30.33$). The difference between the answers to the items related to the velar [k] and bilabial [p] remained inconsistent among items of both sections of the questionnaire.

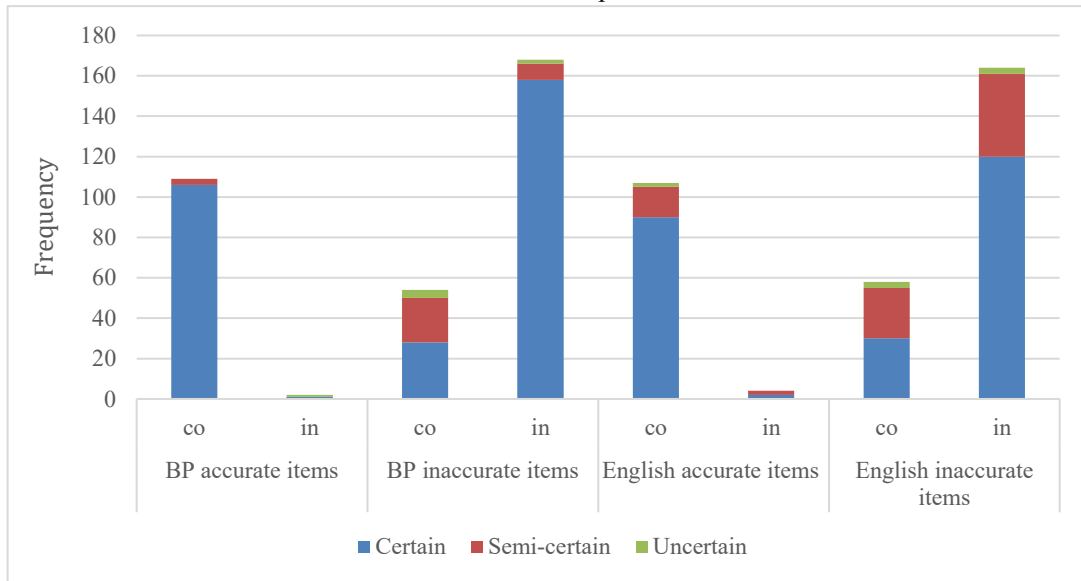
Table 11 – Frequency and relative frequency (in percentage) of the certainty ratings for each item type according to participants' correct and incorrect responses and for all responses together

	BP accurate items (n=111)			BP inaccurate items (n=222)		
	correct	incorrect	all	correct	incorrect	all
Certain	106 (95.5%)	1 (0.9%)	107 (96.4%)	28 (12.61%)	158 (71.17%)	186 (83.78%)
Semi-certain	3 (2.7%)	0 (0%)	3 (2.7%)	22 (9.91%)	8 (3.6%)	30 (13.51%)
Uncertain	0 (0%)	1 (0.9%)	1 (0.9%)	4 (1.8%)	2 (0.9%)	6 (2.7%)
	English accurate items (n=111)			English inaccurate items (n=222)		
	correct	incorrect	all	correct	incorrect	all
Certain	90 (81.01%)	2 (1.8%)	92 (82.88%)	30 (13.51%)	120 (54.05%)	150 (67.57%)
Semi-certain	15 (13.51%)	2 (1.8%)	17 (15.32%)	25 (11.26%)	41 (18.47%)	66 (29.73%)
Uncertain	2 (1.8%)	0 (0%)	2 (1.8%)	3 (1.35%)	3 (1.35%)	6 (2.7%)

Source: the author.

One of the main differences between the *Error Detection Test* and the *Audio-based section* of the questionnaire measuring explicit awareness was that the latter included certainty ratings (Certain – Semi-certain – Uncertain), aiming to verify whether participants' answers were mostly based on intuition or explicit awareness. Table 11 displays the frequency and relative frequency (in percentage) of certainty ratings for each item type, according to participants' correct and incorrect responses, and for the responses together. Figure 16 graphically summarizes the information. It is possible to see that participants were mostly certain of their answers, except for their correct answers to the BP items and to the English items with inaccurate VOT, for which they also indicated being 'semi-certain' for a relatively high number of items. The 'uncertain' rating was rarely used among the item types, with higher relative frequencies (1.8%) for the correct answers to the BP inaccurate and English accurate items. For the audio item type included in the computation of the *Explicit Awareness Score*, namely the items in English with inaccurate VOT, participants were 67.57% certain of their answers, of which 80% were incorrect, 29.73% semi-certain of their answers, of which 62.12% were incorrect, and only 2.7% uncertain of their answers, of which 50% were incorrect. Therefore, if the 'certain' rating resembles more explicit awareness employed in answering the questions, as participants are believed to know the knowledge they are using, participants were likely using the intended type of knowledge, given the contrast in relative frequency of the 'certain' (67.57%) and 'uncertain' (2.7%) ratings used.

Figure 16 – Frequency of the certainty ratings for each item type according to participants’ correct and incorrect responses.



Source: the author.

Key: co = correct responses; in = incorrect responses; BP = Brazilian Portuguese.

Finally, the open-ended responses from the *Text-based section* of the questionnaire measuring explicit awareness may elucidate what participants considered when answering the statements about stops in initial position. Table 12 provides the open-ended responses per statement for the first set of statements and per group of statements for the second set, as the statements for [t] and [k] were grouped on the questionnaire. The first thing that stands out when looking at the table is the notable difference in the number of responses among the first set of statements: 24 responses to [t], 5 to [k], and only 2 to [p]²³. This resembles the unbalanced number of ‘no’ (correct) answers to the corresponding items, 26 to the alveolar [t], 6 to the velar [k], and 2 to the bilabial [p]. Although the number of responses related to each stop in the second set of statements is not as clear as for the first set, since those were grouped into three statements, 7 responses somehow mention the alveolar [t] (V – p02, p08, p09, p11, p16, p20, p32), 4 responses seem to be related to the velar [k] (IV – p02, p10, p19, p32), while only one response is more directly related to the bilabial [p] (VI – p11).

²³ Future research could investigate the participants’ BP variety as palatalization of the alveolar [t] is common in different BP varieties, and this could interfere with the perception of English VOT of this voiceless stop.

Table 12 – Open-ended responses to the first and second sets of statements related to stops from the *Text-based* section of the questionnaire measuring explicit awareness

FIRST SET OF STATEMENTS - Você concorda com as seguintes frases? (não)	
I	1. A pronúncia do 'T' na palavra em inglês 'TO' (ex.: go <u>t</u>o bed) é igual à pronúncia do 'T' na palavra em português 'TU' (ex.: <u>t</u>u és linda).
p01	o som em português é mais longo
p02	No som em inglês, a impressão é que um certo assobio se forma, já no português, é um som fechado, como se fosse uma imitação de uma percussão, por exemplo.
p03	em "go to bed" o som do T é similar à pronúncia de "two"
p04	na palavra "to" o T tem um som mais chiado, como se houvesse um TH. Já em português, na palavra "tu" o T tem um som mais forte e a língua bate nos dentes para pronunciar.
p06	to - é como se o "t" estivesse junto com algo parecido com o "s", só que mais leve. "ts-o" tu - mais seco, só o som do "t" e do "u"
p07	o som em "to" é mais "chiado"
p08	em TO o som é mais tchu enquanto em TU é o som de como se escreve
p11	no inglês eu pronuncio o 'T' um pouco arrastado, diferentemente de como faço para pronunciar 'TU' em português. No entanto não sei se seria necessariamente errado pronunciar da mesma forma
p12	TCHU/TU
p13	o t em "to" lembra o som em português para o r (fonema vibrante simples, r "brando") embora não seja igual
p16	em português é "tu" e em inglês geralmente é "ru"
p17	To - too
p19	o "t" em inglês é mais suave
p20	tsho / Tu
p22	um é tipo tchu e o outro tu
p23	O T em inglês tem o som de T+U ou W seguido do som do O. O T em português tem o som seco do fonema T seguido do fonema U.
p25	O som do "to" em inglês é mais sutil do que o "tu" em português, parece que a boca articula menos para falar "to" do que em "tu".
p26	A pronúncia do T entre vogais em inglês tende a ter o mesmo som de "r" entre vogais em português. Mas existem variações linguísticas. Provavelmente em algum local o som do T na palavra "to" se aproxima do T na palavra "tu".
p28	"to" softer touch on palate than "tu"
p29	Tu - T com língua nos dentes

To - Ts? língua no palato?

p31 A pronúncia em TU do português parece ser mais "fechada" ou "rípida"

p32 Em inglês a pronúncia do "to" é "tsu". Duas letras, três sons.
Em fato nessa sentença corrida o "to" sai "ru".
Em português o "tu" tem duas letras e dois sons.

p35 A palavra "to" tem som de tch - o, enquanto a palavra "tu" tem som de t - u.

p37 O T na palavra To em inglês é pronunciado de forma mais suave, seguindo a mesma pronúncia de "two" e "too". Enquanto no português para Tu, é mais definido, tendo o mesmo som de quando falamos "tigre".

Na prática, percebemos que no inglês para pronunciar o T, precisamos encostar a parte de cima da língua no céu na boca. Enquanto no português, encostamos a ponta dela.

II 2. A pronúncia do 'C' na palavra em inglês 'CAR' (ex.: drive a car) é igual à pronúncia do 'C' na palavra em português 'CARRO' (ex.: dirijo um carro).

p02 Em inglês, o som parece mais fechado de certa forma.

p07 som mais fechado

p17 C em inglês similar ao k

p30 Eu acho que a diferença está na entonação de Car e Carro.

p36 Na palavra "Carro" o CA tem som de KA

III 3. A pronúncia do 'P' na palavra em inglês 'PET' (ex.: my pet dog) é igual à pronúncia do 'P' na palavra em português 'PÉ' (ex.: meu pé direito).

p02 Nos dois, é uma sensação de "estouro", mas no inglês parece um estouro mais projetado e mais ar sai.

p28 Dependendo do sotaque, há uma "aspiration" junto do p de Pet

SECOND SET OF STATEMENTS - Você concorda com as seguintes frases? (sim)¹

IV 1. Brasileiros têm um sotaque característico ao pronunciar o 'C' nas palavras em inglês: 'CALL', 'COOL' e 'CAP'.

V 3. Brasileiros têm um sotaque característico ao pronunciar o 'T' nas palavras em inglês: 'TALL', 'TOP' e 'TASK'.

p02 1. No inglês é como se o som se formasse mais pra cima, perto do palato. Já no português é como se fosse mais pra baixo, na glote.
3. No T em inglês sinto um fluxo de ar maior ao pronunciar, enquanto no português o som sai sem muito ar.

p07 No meu ponto de vista, na língua portuguesa, o som dessas letras é representado de uma forma que quando aplicado no inglês não fica correto, em resumo, o brasileiro não está acostumado com o som de algumas letras em inglês, e por isso a apresenta um sotaque

-
- p08** Acredito que, ao pronunciar palavras com T, os brasileiros usam mais a língua, tocando os dentes com ela, enquanto em inglês, o som é mais curto e usa menos a língua. A diferença é muito sutil, e quase imperceptível em uma conversa normal.
-
- p09** Muitas vezes falam o T com algo como um impulso a mais, tipo vai falar tall e falar algo como t-hall, como se estivesse usando a língua tipo no TH do inglês, mas com uma soltada de ar a mais.
-
- p10** Concordo com o 1. O C nessas palavras em inglês parecem ter um som mais longo, os brasileiros falam mais semelhante com o som em português
-
- p11** Creio que brasileiros em geral tendam a não arrastar o "T" tanto quanto o fazem os nativos de língua inglesa
-
- p15** A tendência do brasileiro ao pronunciar uma letra em inglês, é pronunciar-la conforme um padrão já exercitado e aprendido ao longo dos anos de pronúncia da mesma letra em português.
-
- p16** (3) no português parece que colocam mais ênfase na letra que vem após o T, como TÁ e TÓ
-
- p19** 1) Eu acho que há sons variados da letra "c", por exemplo em "cap" é mais suave que em "cool". Já no português, não há essa variedade.
-
- p20** ao pronunciar o T, os falantes de inglês nativos soam como "tsh", já os brasileiros o pronunciam de modo mais incisivo e breve, como na palavra "tu"
-
- p23** Os brasileiros tendem a pronunciar estes fonemas mais parecidos com o que eles soam em português.
-
- p32** A pronúncia do "C" é mais seca quando alguém (iniciante brasileiro) está aprendendo inglês. O mesmo vale para o "T" que em algumas palavras em inglês tem um "S" muito sutil ali no meio.
-
- p35** Acho que o sotaque brasileiro nessas palavras não está ligado ao som da letra inicial e sim à pronúncia do resto da frase.
-

VI 3. Brasileiros têm um sotaque característico ao pronunciar o 'P' nas palavras em inglês: 'PACK', 'POOL' e 'POP'.

- p07** O fonema de tais letras em inglês se difere do fonema em inglês
-
- p11** Talvez, brasileiros em geral tendam a pronunciar o J sem efetuar um som incidental de "D" e o "P" mais seco
-
- p23** Os brasileiros tendem a pronunciar estes fonemas da forma como eles soam em português. As demais letras que vêm depois do fonema inicial também parecem ter impacto no som da palavra como um todo.
-

Source: the author.

¹ The responses that explicitly mentioned the distractor sounds only were not included in this list, whereas responses that referred to sounds in general were included together with the responses that mentioned the target sounds.

In the explanations of how the English and BP stop sounds differ, participants varied widely in the degree of metalanguage; some were more explicit about VOT, while others seemed to be referring to something else. For example, some open-ended responses to the first set of statements related to the alveolar [t] that reasonably mention VOT are from p02 and p07

(I), in which a ‘whistle’ or ‘hissing’ is mentioned. For the velar [k], the explanations seem to be related to things apart from VOT, such as the one from p30 (II), which mentioned intonation; the only response that could possibly be related to VOT is from p02, which mentioned ‘more closed sound’. For the bilabial [p], both responses may be interpreted as related to VOT as they mentioned ‘burst/more air flow’ (p02 – III) and ‘aspiration’ (p28 – III). However, some responses explicitly mention other things that are not VOT, such as the ones from p13 – I and p16 – I, which mentioned alveolar flapping, probably because of the example given for the English [t] sound, which was in an intervocalic position (go to bed). Therefore, the questionnaire aimed at assessing VOT explicit awareness proved to do that to some extent, although some participants appeared to be answering the questions with other phonetic-phonological aspects in mind.

4.1.1 Discussion of RQ1

To answer RQ1, data from the instruments measuring implicit and explicit phonological awareness about English VOT were analyzed. For implicit phonological awareness, data from the *Error Detection Test* indicate that Brazilian learners of English possess low levels of implicit phonological awareness about English VOT. The mean *Implicit Awareness Score* was slightly above the chance level (Figure 13), but when looking into the percentages of correct responses to the accurate and inaccurate tokens (Figure 12), the inaccurate tokens were mostly identified as correct, just like the accurate tokens, which suggests that participants considered both types of tokens to be correct, pressing the ‘correct’ answer for most words. This difficulty in identifying inaccurate tokens as incorrect shows a lack of implicit phonological awareness, as this type of knowledge is frequently associated with the ability to identify deviations in speech (Ellis, 2005). Kivistö-de Souza (2017) had already observed this difficulty by Brazilian learners of English in identifying inaccurate English VOT as incorrect, so data from the present study confirm the extremely low levels of implicit awareness about this English phonetic feature by the target population.

As shown by studies reviewed previously, Brazilians do not use VOT as the main acoustic cue to differentiate voiced from voiceless stops in either BP or English; instead, they seem to rely on other cues, such as burst intensity or f_0 , to distinguish the two sets of stops (Alves; Kampff, 2019; Alvez; Zimmer, 2015). This may have been the underlying reason participants tended to indicate that both tokens with accurate and inaccurate VOT values were correct once the manipulated inaccurate tokens maintained the original acoustic cues from the

English voiceless stops, except for the VOT that was cut off to create tokens with Zero VOT. Thus, the voiceless stops with Positive and Zero VOT possibly sounded correct to Brazilian learners who were judging accuracy based on other acoustic cues. However, as Alvez and Zimmer (2015) suggested, at least to a limited extent, Brazilians seem to recognize that English voiceless stops are produced with longer VOT values than those in BP, as they already produce English voiceless stops with longer VOT values than their BP values since the first levels of proficiency. Nevertheless, participants from the present study considered most voiceless stops of the *Error Detection Test* to be correct, regardless of the VOT pattern (Positive or Zero VOT).

For explicit phonological awareness, data from the *Audio-based* and *Text-based sections* of the questionnaire measuring this type of knowledge demonstrate that Brazilian learners possess extremely low levels of explicit phonological awareness about English VOT. The low mean *Explicit Awareness Score* ($M = 30.63$) is the main indication of the lack of explicit phonological awareness. Moreover, looking into the first set of statements of the *Text-based section*, very few participants agreed that English and BP voiceless stops differ in the initial position of monosyllabic words, except the alveolar [t], which more participants agreed to be different when comparing the two languages – although the open-ended responses showed that not all the agreements were considering VOT, but some were taking other phonetic aspects into account. Still, a considerable number of participants mentioned something that can be interpreted as VOT in their open-ended responses for [t]. Therefore, the Brazilian learners of English may differ in their levels of explicit awareness of English VOT among the stops. If this is the case, analyses like this should look for the differences in VOT awareness among stops separately and not only take VOT as a single construct.

In addition to that, data from the *Audio-based section* of the questionnaire measuring explicit awareness appear to confirm that Brazilians were using other acoustic cues apart from VOT to judge the accuracy of voiceless stops in the instruments of the present study. This is because Brazilians showed difficulty in identifying inaccurate items in BP as incorrect, showing that VOT, the only acoustic cue manipulated in the stimuli, was not being used to judge the accuracy of the voiceless stops. Therefore, participants likely applied the rule governing stops of their L1 – in which VOT plays a marginal role, and thus those audio-inaccurate items in BP that had their VOT manipulated were considered accurate by participants – to what is accurate in the L2. In applying L1 rules to the L2, VOT was not relevant for the meaning, maybe taken as an ‘allophonic’ variation by participants, which led to considering most tokens correct despite the VOT pattern.

Therefore, data from the instruments measuring implicit and explicit awareness supported the hypothesis for RQ1 that Brazilian learners of English possess low degrees of implicit and explicit awareness about English VOT. Participants demonstrated even lower degrees of explicit awareness than of implicit awareness, which is in line with previous studies suggesting that L2 phonological awareness consists mostly of implicit awareness (Kivistö-de Souza, 2015).

4.2 RQ2: THE RELATION BETWEEN IMPLICIT AND EXPLICIT KNOWLEDGE

Research Question 2 asked whether there is a relation between implicit and explicit L2 phonological awareness. The hypothesis is that, given the theoretical foundations of the weak-interface position (Ellis, 2006; Schmidt, 1995) and the empirical findings that implicit and explicit L2 morphosyntactic knowledge are reciprocally associated (Kim; Godfroid, 2023; Suzuki; DeKeyser, 2017), there will be a relation between implicit and explicit L2 phonological awareness. To answer this question, a correlation was conducted between the scores of implicit and explicit awareness.

The examination of the correlation assumptions indicated that the *Implicit Awareness Score*²⁴ did not fit a normal distribution, as determined by the *p-value* of the Shapiro-Wilk test ($p < .001$). Therefore, a non-parametric Spearman's correlation between the *Implicit Awareness Score* and the *Explicit Awareness Score* was run. The analysis revealed a positive but non-significant relationship between implicit and explicit knowledge, $r_s = .26$, $p = .12$, 95% CI $[-.07, .54]$, $R^2 = .06$, $N = 37$. Although the correlation did not reach significance, the effect size was close to medium, as indicated by the correlation coefficient itself ($r_s = .26$), and the two variables had a small-to-medium shared variance of approximately 6%, as expressed by the squared r_s ($R^2 = .06$).

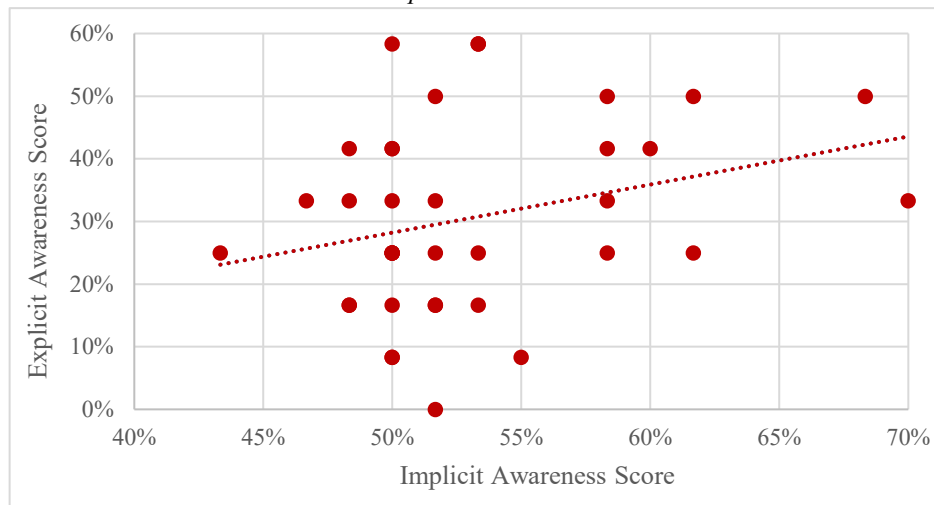
The power test for the correlation was computed in RStudio using the 'pwr.r.test()' function from the 'pwr' package following Larson-Hall (2015, p. 106). The significance level and type of hypothesis test were set to the default of .05 and two-tailed, respectively. The correlation power level was .35, meaning there was only a 35% chance of finding statistical

²⁴ In the present study, implicit awareness was taken as the ability to identify accurate and inaccurate samples of the target form (VOT), and that is why the *Implicit Awareness Score* was used here. However, as the results for RQ1 show, Brazilians seem to have considerably more difficulty in identifying inaccurate samples as incorrect than in identifying accurate samples as correct, at least when it comes to VOT. Thereupon, further research should consider other ways of operationalizing implicit awareness, perhaps analyzing it as the ability to identify pronunciation deviations only. It is important to note that a correlation with the percentage of correct responses to the incorrect tokens only was not run, and results could have been different if that were the case.

significance, and a 65% chance of not finding it, given the sample size ($N = 37$) and the effect size (.26). To reach the adequate power of .80, proposed by Murphy and Myors (2023), with the same effect size (.26) it would be necessary to increase the sample size to 113 participants. If the type of hypothesis test was set to a one-tailed test for a positive correlation, the number of participants needed for a power level of .80 would decrease to 89, with the same effect size of .26. However, the effect size of .26 is closer to the medium magnitude of .30 proposed by Cohen (1992) than to the small magnitude of .10, and if we suppose there is a weaker relationship between implicit and explicit awareness than the one encountered here (.26) – although lacking significance –, the number of participants needed to reach the proposed power of .80 would increase substantially. For a two-tailed test with a small effect size of .10, it would be necessary to include 782 participants in the analysis to reach the power of .80. For the same power and effect size, but in a one-tailed test in the direction of a positive correlation, it would be necessary to include 616 participants. It should be noted that the ‘pwr.r.test()’ function was originally designed to compute the power of Pearson’s correlations, so interpretations must be done with caution.

For visual aid, Figure 17 shows a scatter plot of the correlation, displaying the *Implicit Awareness Score* on the horizontal axis and the *Explicit Awareness Score* on the vertical axis. A line of the best fit was included to illustrate the overall trend in the data. The plot area was targeted to where the dots are distributed by adjusting the range of values of the axes to assist the analysis. Therefore, the vertical axis represents the lower end of the scale for explicit awareness (0% - 60%), while the horizontal axis indicates the middle of the scale for implicit awareness (40% - 70%). The line of best fit has a slight positive slope, indicating a weak positive correlation between the variables, i.e., higher explicit awareness tends to be associated with higher implicit awareness, although this relationship still lacks statistical significance. By looking into the data points, the two dots closest to the right side of the plot are above the closest dot to the left side of the plot, following the line of best fit. However, the dots in between those in the side borders show no apparent pattern between the two variables, which may partially explain, together with the relatively small sample size ($N = 37$), the failure to find a statistically significant relation.

Figure 17 – Scatter plot of the correlation between the *Implicit Awareness Score* and the *Explicit Awareness Score*



Source: the author.

4.2.1 Discussion of RQ2

To answer RQ2, the correlation between the implicit and explicit awareness was analyzed. Although the correlation was not statistically significant, the correlation coefficient and related statistics yielded interesting results. As for the lack of statistical significance, several possible explanations may justify it. One major reason is the limited statistical power of the correlation.

As argued by Murphy and Myers (2023), the likelihood of finding statistically significant results in a study is very susceptible to its statistical power level, which, in turn, is directly impacted by sample size. Therefore, other parameters apart from the p -value, such as the effect size – considered by Murphy and Myers (2023, p. 106) to be “the most important parameter of all” – may yield more important information about the variables regardless of the statistical significance. Given the little power of the correlation between implicit and explicit awareness (.35), with only a 35% chance of finding a real effect (Larson-Hall, 2015), the lack of statistical significance was somewhat expected, which classifies the present study as exploratory in nature. Therefore, looking at the effect size of .26, the strength of the correlation between implicit and explicit awareness is close to the medium magnitude value established by Cohen (1992) of .30. A moderate effect size can be considered of satisfactory magnitude in social and behavioral science, as effect sizes in this field tend to be small to moderate (Murphy; Myers, 2023). Moreover, the squared correlation coefficient ($R^2 = .06$) expressed a small-to-medium shared variance between the two types of knowledge, approximately 6%, which can be considered sufficiently satisfactory in the field. Finally, the 95% confidence interval [-.07,

.54] suggested that if the study were replicated 100 times, 95% of the time the correlation coefficient would be found within the range from -.07 to .54. As an exploratory study, this advises further research that there may be a tendency of the correlation between implicit and explicit L2 phonological awareness to be positive, i.e., higher implicit awareness is related to higher explicit awareness.

The present study is exploratory in nature, not only due to the limited power impacted by the relatively small sample size, but also due to the lack of validity of the novel questionnaire intended to measure explicit L2 phonological awareness. Another possible explanation for the lack of statistical significance regards the *Explicit Awareness Score* used in the correlation, which is derived from the *Audio-based* and *Text-based sections* of the questionnaire measuring explicit awareness. As this is a novel questionnaire developed in the present study, time constraints imposed by external factors hindered the possibility of conducting pilot studies and refining the data analysis to create a composite score for explicit awareness, as no similar instrument was found in the literature. Therefore, the questionnaire may not fully capture what it is intended to, and the *Explicit Awareness Score* may not be the best representation of the variable.

Moreover, the operationalization of implicit awareness in the present study may also have impacted the findings. Implicit awareness was operationalized in the present study as the ability to identify accurate and inaccurate samples of VOT. However, as demonstrated by the results for RQ1, participants had substantially more difficulty in identifying inaccurate tokens as incorrect than accurate tokens as correct. Therefore, the difficulty in identifying accurate and inaccurate samples of VOT is not the same, suggesting another possible operationalization of implicit awareness. Implicit awareness could be analyzed as the ability to detect pronunciation deviation in inaccurate tokens, excluding the ability to accept accurate tokens as correct. This could result in different findings and methods. Nonetheless, as indicated by the effect size, the use of the *Implicit Awareness Score* and the *Explicit Awareness Score* in the correlation yielded satisfactory strength and can be considered adequate given the exploratory nature of the study.

As a result, the lack of statistical significance in the correlation between the two types of knowledge failed to empirically support the hypothesis for RQ2 that there is a relation between implicit and explicit L2 phonological awareness. However, the limited statistical power and other statistical results, such as the effect size, indicated the possibility of a Type II error, which is when the test fails to detect a real effect.

4.3 RQ3: VARIANCE IN IMPLICIT L2 AWARENESS ABOUT VOT

Research Question 3 asked how much of the variance in implicit L2 phonological awareness about VOT can be explained by L2 proficiency, L2 input, and explicit L2 phonological awareness. The hypothesis was that, in accordance with the revised Speech Learning Model (Flege; Bohn, 2021) that posits that L2 input affects L2 perception, and based on findings from Kivistö-de Souza (2015), Kim and Godfroid (2023) and Suzuki and DeKeyser (2017), a considerable amount of implicit L2 phonological awareness will be explained by L2 proficiency, L2 input, and explicit L2 phonological awareness. To answer RQ3, a multiple regression analysis was conducted with implicit knowledge as the dependent variable, and L2 proficiency, L2 input, and explicit L2 phonological awareness as independent variables.

When examining the assumptions for running a multiple regression, the presence of an outlier was verified by finding one standardized residual value larger than 3 (3.306). This participant (p24) was removed from the multiple regression analysis, which was then rerun with the remaining 36 participants. Thus, the following results do not include this outlier. Cook's distance values were also checked for outliers and no unusually large values were found, with the maximum distance being .311. This indicated the absence of other outliers in the analysis.

To check for multicollinearity and relation, multiple correlations between the variables were run (see Table 13). Spearman's correlations were conducted with the *Implicit Awareness Score* because it did not follow a normal distribution. The other three correlations were the parametric Pearson's correlations. The correlations between the dependent and independent variables were weak and positive, although non-significant. Regarding multicollinearity, the independent variables were not highly intercorrelated, although they also lacked significance (Table 13). At last, the absence of multicollinearity was verified with tolerance values above 0.2, and Variance Inflation Factor (VIF) values below 5 (Table 14).

Table 13 – Correlation coefficients between the variables ($N = 36$)

Variable	(1) ¹	(2)	(3)	(4)
(1) Implicit awareness	—			
(2) L2 proficiency	.20 ($p.24$)	—		
(3) L2 input	.02 ($p.93$)	.27 ($p.12$)	—	
(4) Explicit awareness	.26 ($p.13$)	-.26 ($p.13$)	-.31 ($p.07$)	—

Source: the author.

¹ The correlations with implicit awareness are Spearman's correlations, as the *Implicit Awareness Score* did not fit a normal distribution. The other three correlations are Pearson's correlations.

Linearity and homoscedasticity of data were visually inspected through the residuals versus predicted values plot, which showed no clear pattern or funnel shape that violates assumptions. The Q-Q plots of standardized residuals showed a reasonably normal distribution, with only minor deviations at the tails. Moreover, it is known that multiple regressions are susceptible to the sample size, although there is no agreement on the recommendations of the sample size needed per independent variable (Larson-Hall, 2015). With the current three independent variables, Cohen *et al.* (2003) would recommend 26 participants to have 80% power in finding a moderate effect and 91 participants in finding a small effect (Larson-Hall, 2015). With the existing sample size ($N = 36$), there is limited power to find a small effect, although it is easier to find larger effects or even a small one, considering other recommendations cited in Larson-Hall (2015) that suggest at least 10 participants for each independent variable. Therefore, preliminary analysis showed assumptions were reasonably met.

Table 14 – Results from the multiple regression analysis ($N = 36$)

Predictor	<i>B</i>	<i>SE B</i>	β	<i>t</i>	<i>p</i>	95% <i>CI</i>		Collinearity	
						Lower	Upper	Tolerance	<i>VIF</i>
Intercept	37.278	6.12	-	6.093	< .001	24.816	49.739	-	-
L2 proficiency	.141	.068	.343	2.069	.047	.002	.281	.893	1.12
L2 input	.025	.095	.044	.259	.797	-.17	.219	.866	1.16
Explicit Awareness	.135	.054	.417	2.485	.018	.024	.246	.871	1.15

Source: the author.

Key: *SE* = Standard Error; *CI* = Confidence Interval; *VIF* = Variance Inflation Factor.

Hence, a standard multiple regression analysis was run with implicit awareness as the dependent variable and explicit awareness, L2 input, and L2 proficiency as independent variables. Together, the independent variables explained about 22% of the variation in implicit awareness scores ($R^2 = .216$, $F(3, 32) = 2.936$, $p.048$). Out of the three independent variables, explicit awareness had the greatest impact on implicit awareness ($\beta = .417$, $t(32) = 2.485$, $p.018$), followed by L2 proficiency ($\beta = .343$, $t(32) = 2.069$, $p.047$). In contrast, L2 input ($\beta = .044$, $t(32) = 0.259$, $p.797$) did not significantly impact implicit awareness scores. When analyzing the unique contribution of each variable to the implicit awareness scores by obtaining the squared semipartial correlations (rs^2), explicit awareness accounted for around 15% of the variance ($rs^2 = .15$), while L2 proficiency accounted for around 10% ($rs^2 = .10$). Moreover, the intercept value (β_0), that is, the value of the dependent variable when the three independent variables are at zero, was 37.278 ($SE \beta_0 = 6.12$, $t(32) = 6.093$, $p.<.001$).

4.3.1 Discussion of RQ3

To answer RQ3, results from the multiple regression were interpreted. The model explained around 22% of the variance in implicit L2 phonological awareness.

Interestingly, explicit L2 phonological awareness and L2 proficiency, but not L2 input, were found to significantly predict implicit L2 phonological awareness. Explicit awareness uniquely accounted for most of the variance (15%) in implicit awareness, accounting for a stronger unique contribution than L2 proficiency (10%). This indicates that implicit awareness about English VOT tends to be somewhat impacted by explicit awareness of the same phonological feature, as suggested by the weak-interface position (Ellis, 1994). Although to a lower degree, L2 proficiency was also shown to impact implicit awareness. One possible explanation for the contribution of proficiency is VanPatten's (1990) postulation of the primacy of meaning over form, in which language learners attend to language form only when attentional resources have not been depleted by attending to meaning. Although VanPatten (1990) refers more directly only to grammatical morphology, this postulation may be extended to phonetics and phonology. In this way, more proficient learners may possess more attentional resources available to attend to formal characteristics of L2 speech than lower-proficient learners, whose attention is being used to decipher meaning.

As for the lack of association between L2 input and implicit L2 phonological awareness, this might reflect a failure in the measurement or in the operationalization of the construct. In relation to the measurement of L2 input, the current study employed a questionnaire about the contexts in which participants had contact (in hours) with the L2. However, studies in instructed language learning settings have incorporated other measurements. Saito and Hanzawa (2018), for example, apart from including information about the L2 use outside the classroom, also encompassed the length of instruction (the total hours of English classes) and the type of instruction (form- vs. meaning-oriented classes) in their operationalization of L2 input, which was found to have a significant impact on some areas of oral proficiency development. Moreover, as previously cited, more sophisticated measurements of L2 input that go beyond learners' self-reports on quantitative scales are demanded in the area (Mayor, 2009). In relation to the operationalization, L2 input was operationalized here as one single construct, and no separation between quantity and quality of L2 input was made in the analysis. As demonstrated in previous research, L2 quality but not quantity appears to play a major role in L2 speech development of segments in instructed language learning settings (Saito; Hanzawa, 2018; Teló;

Kivistö-de Souza, 2023). Therefore, not looking for the impact of L2 quality and quantity separately is a limitation of the present study. No separation was done due to the relatively small sample size, and so adding one more independent variable to the multiple regression analysis would potentially lower the power of the model.

Therefore, the model provided empirical support for part of the hypothesis for RQ3. Two out of the three variables were found to predict implicit L2 phonological awareness about VOT: L2 proficiency and explicit L2 phonological awareness, but not L2 input.

4.4 GENERAL DISCUSSION

The present study aimed to measure the implicit and explicit awareness of English VOT by Brazilian learners of English through a novel questionnaire developed to assess explicit phonological awareness and a perception test to assess implicit phonological awareness available in the literature (Flege; MacKay, 2004). Due to the innovative nature of the questionnaire as well as the considerably limited sample size of the study, the results are merely tentative in the sense that they uncover a possibility to assess what has not been empirically tested, to the best of our knowledge, that is, explicit L2 phonological awareness. Nonetheless, results provide implications for future research by shedding light on the long-standing discussion on the possible interface between implicit and explicit L2 knowledge.

In terms of the levels of implicit and explicit awareness of English VOT, Brazilian learners of English appeared to possess extremely low levels of both types of knowledge. In the *Error Detection Test*, measuring implicit awareness, participants scored slightly above the chance level ($M = 53.15$, $SD = 5.7$, range = 43.33 – 70), with most answers indicating that the English voiceless stops were correct, regardless of the accuracy of their VOT values. This corroborates the difficulty observed in Kivistö-de Souza (2017) by this population in identifying incorrect English VOT. In line with Kivistö-de Souza (2015), who suggested that L2 phonological awareness consists mostly of implicit awareness, the participants in the present study scored even lower in the questionnaire measuring explicit awareness ($M = 30.63$, $SD = 15.22$, range = 0 – 58.33) than in the *Error Detection Test*. In the open-ended questions, the lack of explicit knowledge about English VOT became clearer as few to no participants referred to this English phonological aspect in their responses, except for the alveolar [t], in which Brazilians seem to be somewhat more aware of its aspiration in English.

Regarding the relationship between implicit and explicit knowledge and the factors affecting implicit L2 phonological awareness about VOT, the results provide interesting insights into the interface hypotheses. Although the correlation between the implicit and explicit knowledge did not reach significance, as expected due to the limited power (35%), the effect size was close to moderate (.26). Moreover, in the multiple regression, explicit awareness was found to be the stronger predictor of implicit awareness in the model, uniquely accounting for approximately 15% of variance. Therefore, there seems to be a relationship between the implicit and explicit L2 knowledge, which empirically supports the postulations of the weak-interface hypothesis (Ellis, 1994). However, further studies with larger sample sizes and more advanced measurements of L2 knowledge should be conducted to confirm the trend observed here. As

for the other two variables analyzed to impact implicit awareness of English VOT, only L2 proficiency, but not L2 input, was found to be a significant predictor, uniquely accounting for approximately 10% of the variance in the *Implicit Awareness Score*. The effect of L2 input on L2 phonological awareness remains empirically under-researched, primarily in instructed language learning settings. Previous research has suggested that L2 input quality may better explain the variation in the development of L2 segments in this language learning setting than L2 input quantity (Teló; Kivistö-de Souza, 2023; Saito; Hanzawa, 2018), which may partially explain the absence of impact found here, given that the operationalization of L2 input consisted mostly of L2 input quantity. In general, more empirical investigations are necessary to explore the nature of implicit and explicit L2 knowledge and the factors affecting them.

5 CONCLUSIONS

There is no clear consensus in the field of Second Language Acquisition regarding the definitions and operationalizations of implicit and explicit L2 knowledge, as well as the relationship between these two types of knowledge. The three interface hypotheses, extensively addressed in theoretical studies, propose three main views on the interface between implicit and explicit L2 knowledge: the *non-interface*, the *strong interface*, and the *weak interface*. The possible relationship between these two types of knowledge is of extreme importance in the field of SLA given the nature of explicit instruction that mostly results in explicit knowledge, which in turn, is discussed to impact or not on the development of implicit knowledge – the subconscious type of knowledge used in efficient communication (Ellis, 2005). Despite the extensive theoretical discussion on the interface hypotheses, empirical investigations of the matter are hindered by the lack of consensus on the definitions and operationalizations of implicit and explicit L2 knowledge, as well as by the lack of instruments available in the literature to measure such constructs (Roehr-Brackin, 2024).

The present thesis empirically investigated implicit and explicit L2 awareness, their relationship, and the factors affecting implicit L2 awareness. The study was conducted with 37 Brazilian learners of English who were tested on their implicit and explicit awareness of English VOT through an online experiment containing four instruments to measure the two types of knowledge, as well as L2 proficiency and L2 input, which were variables analyzed to impact implicit L2 awareness, together with explicit L2 awareness.

The exploratory results indicate that Brazilian learners of English possess generally low levels of implicit and explicit awareness of English VOT. This was expected considering the similar results from previous research (Kivistö-de Souza, 2017) and the different status of VOT assumed in English and Brazilian Portuguese. As VOT is the most important phonological feature in differentiating voiced from voiceless stops in English but not in BP, in which other features like *f0* are used to distinguish [p, t, k] from [b, d, g] (Alves; Kampff, 2019; Kupske; Perozzo, 2024), Brazilians seem to not be aware of the English VOT patterns employed by native speakers of American English. More specifically, Brazilian learners appear to display low levels of implicit and explicit awareness of the English VOT of voiceless stops, the positive VOT value not usually found in BP.

Concerning the relation between implicit and explicit L2 phonological awareness, the tentative results indicate a weak interface between them. The effect size of the correlation between the two types of L2 awareness, although not reaching statistical significance, was close

to moderate. The lack of statistical significance is likely related to the limited power of the correlation due to the relatively small sample size, less than one-third of the sample size needed for adequate power in exploratory studies proposed by Murphy and Myers (2023) with the same effect size. Moreover, in the multiple regression analysis to verify the variables that explain the variance in implicit L2 awareness, explicit L2 awareness was found to be the strongest predictor, accounting for a unique variance of around 15%. Therefore, the results from the present study suggest that implicit and explicit L2 phonological awareness may be somewhat related.

In terms of the variables analyzed to explain the variance in implicit L2 phonological awareness about VOT, explicit L2 phonological awareness about VOT and L2 proficiency, but not L2 input, were found to be significant predictors in the multiple regression analysis. The contribution of L2 proficiency to the variance in implicit awareness may be partially explained by VanPatten's (1990) postulation of 'primacy over meaning', in which more advanced Brazilian learners of English may possess more attentional resources available to attend to formal characteristics of L2 speech once they free those resources from attending to meaning, something that less proficient learners may still be depleting attention in. The lack of contribution of L2 input observed may be due to the measurement or operationalization of the construct used in this study. L2 input was operationalized as time spent in activities in English in a typical week. Other studies in the field have employed a variety of methods to operationalize L2 input, including years and type of instruction, for example, as well as operationalized L2 input according to L2 input *quantity* and *quality*. Moreover, other measurements apart from self-reports on quantitative scales have been proposed as better estimations of L2 input (Mayor, 2009). However, few studies have analyzed the influence of L2 input on implicit L2 phonological awareness, and results are still inconclusive. Results from the present study suggest that L2 input, when operationalized mostly in regard to L2 quantity and measured on self-reported scales, does not play a major role in implicit L2 phonological awareness in an instructed language learning setting. Nevertheless, more studies are necessary to draw more concrete conclusions on the nature of implicit and explicit L2 phonological awareness, their relation, and the factors affecting their development, especially studies analyzing the implicit and explicit L2 phonological awareness of a more prominent phonetic-phonological feature than VOT, for which participants are likely to possess some awareness, so the analyses can better predict the relationship between those existing types of awareness and other independent variables.

5.1 LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

As with any other scientific study, the present thesis is subject to several limitations that limit the interpretation and generalization of results. The first substantial limitation of the study involves the novel instrument developed here intended to assess Brazilian learners' explicit awareness of English VOT. Although the creation of this instrument can be considered one of the main contributions of the study, it also presents limitations. As no instrument to measure explicit L2 phonological awareness was found in the literature, to measure explicit L2 awareness of VOT or of any other segment, the instrument developed here had to rely solely on theory. Moreover, due to the limited scope of the thesis and the time constraints imposed, piloting the novel instrument was not possible. Therefore, more research is necessary to conduct extensive piloting using the findings obtained here in order to improve the structure of the questionnaire measuring explicit L2 awareness and the interpretation of the data obtained from it. Future research could include interviews with participants to help elucidate how the instruments are being interpreted by learners and to what extent the instruments are accurately assessing the underlying constructs. Moreover, studies could look into the tests assessing implicit and explicit awareness to see whether the tests themselves can raise awareness of the features analyzed.

The second limitation involves the target phonetic-phonological cue analyzed, the VOT. As participants showed extremely low levels of implicit and explicit awareness about English VOT, the findings are weak. If the phonetic-phonological feature analyzed was more prominent, participants would possibly display higher levels of implicit and explicit knowledge, and the relation between the two types of knowledge, and the contributions of other variables on implicit knowledge could be better analyzed. It is difficult to verify if one type of knowledge that is mostly nonexistent is related to another type of knowledge or to other variables. Therefore, future research should consider investigating easier sounds for learners, such as the sound used as a distractor in the present study, which participants demonstrated to be highly aware of.

The third limitation regards the inclusion of native speakers of American English to perform the instruments. It would be interesting to include native English speakers as participants along with learners of English to provide baseline data on the instruments, mainly for the instruments measuring implicit and explicit L2 phonological awareness.

Another limitation relates to the long-term demand for longitudinal studies, which can help comprehend the development of an L2 with its intricacies, rather than simplifying the

learning process to a single moment. Despite the expressed need, longitudinal studies are particularly difficult to conduct in the Brazilian context due to the legal challenges involved in recruitment, participation, and retention of voluntary participants. Nonetheless, they are pivotal to the understanding of L2 learning, as longitudinal data can elucidate trends not visible in cross-sectional data, allowing for stronger inferences about causality and directionality, such as how L2 input changes over time impacts the degrees of implicit and explicit L2 knowledge. Moreover, experimental studies with at least a short intervention can demonstrate how L2 explicit and implicit phonological awareness is affected by that, and how these two types of knowledge are related to one another, given that intervention usually entails pronunciation instruction, which generally involves some sort of explicit instruction that is believed to develop mainly explicit knowledge. However, analyzing implicit and explicit knowledge over time requires the consolidation of reliable instruments to do so, since the mere application of the instruments used here to tap into those types of knowledge may impact participants' levels of knowledge, and the repeated application may replicate a perceptual training and explicitly train participants on the underlying knowledge intended to be assessed.

Finally, one major limitation of the present study is the relatively limited sample size. Whereas 37 participants can be considered enough, considering the aforementioned challenges in recruiting participants in the context of the current study and the average sample size encountered in similar research, the power level of the correlation and the multiple regression analyses proved to be substantially lower than expected due to the sample size and the exploratory nature of the present study. As is known, in frequentist statistics, statistical significance is somewhat subject to the number of observations and “the use of the p-value imposes a binary and categorical decision” (Lima Jr.; Garcia, 2025, p. 10, our translation). Therefore, future research should aim to include larger sample sizes and consider using Bayesian statistics to minimize the risk of a Type II error, which is the failure to detect a real effect.

5.2 PEDAGOGICAL IMPLICATIONS

The findings of the present study hold significant implications for L2 teaching in general, and more specifically, for English teaching to Brazilian native speakers of Brazilian Portuguese. In terms of the pedagogical implications for L2 teaching in general, the findings from the present study seem to support the weak-interface position between implicit and explicit L2 phonological awareness, which is in line with findings for the relation between implicit and

explicit L2 morphosyntax knowledge (Kim; Godfroid, 2023; Suzuki; DeKeyser, 2017). If the case is that explicit L2 phonological awareness affects the development of implicit L2 phonological awareness, teaching phonology in L2 classrooms can assist language learners in developing the implicit subconscious knowledge used in efficient communication through the development of explicit conscious knowledge in traditional teaching. The findings from this study support that, besides L2 morphosyntax, teaching L2 phonology, which is somewhat less common, may also benefit language learners, given that language teaching is usually related to explicit instruction that is believed to result mostly in explicit L2 knowledge, which in turn, seems to affect the levels of implicit L2 knowledge also known as the “basic linguistic competence” (Ellis, 2004, p. 231).

More specifically, findings have direct implications for English teaching to Brazilian learners of English who were found to possess generally low levels of implicit and explicit L2 phonological awareness about English VOT in the present study. Although in BP, VOT does not constitute a major phonological aspect in differentiating voiced from voiceless stops, in English, VOT is usually the most important phonetic cue for differentiating [p, t, k] from [b, d, g]. This difference between the two languages has implications for native speakers of BP who tend to perceive and produce English voiceless stops with the VOT patterns of their L1. In terms of perception, Brazilians tend to use other acoustic cues than VOT to identify English stops. Regarding production, Brazilians tend to pronounce English voiceless stops with shorter VOT values than those observed in native English speakers, which could result in English voiceless stops with zero, and not positive, VOT, that may be perceived as voiced stops by native English speakers (Alves; Zimmer, 2015). Therefore, the low levels of knowledge about English VOT by Brazilian learners may hinder successful communication in English, given that studies with the same population have shown that L2 phonological awareness contributes to L2 pronunciation (Kivistö-de Souza, 2015; Teló; Kivistö-de Souza, 2023). If language teaching can modify the levels of implicit L2 knowledge as pointed above, Brazilian learners of English must receive instruction about English VOT and its differences from BP VOT. This could be done without mentioning technical terminologies like VOT itself and using a communicative approach, such as the 5-step model proposed by Celce-Murcia *et al.* (2010), successfully used in Kupske and Oliveira (2020).

To conclude, the present study contributed to the research field of Second Language Acquisition, more specifically, this study added to the area of L2 phonological awareness and the heated theoretical debate over the interface hypotheses (Roehr-Brackin, 2024). The focus here was on the implicit and explicit awareness of English VOT, which, unlike BP, consists of

the most important phonetic cue to identify English voiced and voiceless stops. Building on the theory and methods of previous research, a perceptual test was adapted to assess implicit L2 phonological awareness, and a questionnaire was newly designed to assess explicit L2 phonological awareness. The results showed that Brazilian learners of English possessed low implicit and explicit phonological awareness of English VOT. The correlation between these two types of awareness did not reach significance, although the effect size was close to medium (.26). In the multiple regression analysis, explicit L2 phonological awareness uniquely accounted for most of the variance (15%) in implicit L2 awareness, followed by L2 proficiency (10%). Therefore, although the present study advanced the understanding of the relationship between implicit and explicit L2 phonological awareness, more research is needed to draw more concrete conclusions.

REFERENCES

- ALLEN, J.; MILLER, J.; DESTENO, D. Individual talker differences in voice-onset-time. **The Journal of the Acoustical Society of America**, [S. l.], v. 113, n. 1, p. 544–552, Jan. 2003.
- ALVES, M. A. **Estudo dos parâmetros acústicos relacionados à produção das plosivas do português brasileiro na fala adulta: Análise acústico-quantitativa**. Tese (Doutorado em Letras) - Universidade Federal de Santa Catarina, Florianópolis, 2015.
- ALVES, M. A.; SEARA, I. C.; PACHECO, F. S.; KLEIN, S.; SEARA, R. On the voiceless aspirated stops in Brazilian Portuguese. *In*: Computational Processing of the Portuguese Language: 8th International Conference, PROPOR, 8., 2008, Aveiro, Portugal. **Proceedings [...]**. Springer Berlin Heidelberg, p. 248-251, 2008.
- ALVES, U. K. Consciência dos aspectos fonéticos-fonológicos da L2. *In*: LAMPRECHT, Regina Ritter; BLANCO-DUTRA, A. P.; SCHERER, A. P. R.; BARRETO, F. M.; BRISOLARA, L. B.; SANTOS, R. M.; ALVES, U. K. (org.). **Consciência dos sons da língua: subsídios teóricos e práticos para alfabetizadores, fonoaudiólogos e professores de língua inglesa**. Porto Alegre: EDIPUCRS, 2012. Chap. 9, p. 23-26.
- ALVES, U. K. Consciência Fonológica de Línguas Adicionais: Um construto e dois saberes complementares. **Revista signo**, Santa Cruz do Sul, v. 49, n. 95, p. 41–54, May-Aug. 2024.
- ALVES, U. K.; KAMPFF, F. Efeitos de longo prazo do treinamento perceptual na percepção e produção das plosivas iniciais surdas do inglês por estudantes brasileiros: Implicações para o ensino de pronúncia. **Ilha do Desterro**, Florianópolis, v. 72, n. 3, p. 375-400, Oct. 2019.
- ALVES, U. K.; ZIMMER, M. C. Perception and production of English VOT patterns by Brazilian learners: the role of multiple acoustic cues in a DST perspective. **Alfa**, São Paulo, v. 59, n. 1, p. 155-175, Jan-Apr. 2015.
- ANDERSON, J. R. **Acquisition of cognitive skill**. 1. ed. Washington: Psychological Review, 1982.
- ANWYL-IRVINE A. L.; MASSONNIÉ J.; FLITTON A.; KIRKHAM N.; EVERSHERD J.K. Gorilla in our midst: An online behavioral experiment builder. **Behavior research methods**, [S. l.], v. 52, p. 388-407, Feb. 2020.
- ASSOCIATION FOR LANGUAGE AWARENESS. Retrieved from: <https://languageawareness.org/>. Accessed on 2 July 2025.
- AUZOU, P.; OZSANCAK, C.; MORRIS, R. J.; JAN, M.; EUSTACHE, F.; HANNEQUIN, D. Voice onset time in aphasia, apraxia of speech and dysarthria: A review. **Clinical linguistics & phonetics**, [S. l.], v. 14, n. 2, p. 131-150, Jan. 2000.
- BAKER, W.; TROFIMOVICH, P. Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. **Studies in Second Language Acquisition**, [S. l.], v. 28, n. 1, p. 1–30, Mar. 2006.

BARAN, J. A.; LAUFER, M. Z.; DANILOFF, R. Phonological contrastivity in conversation: A comparative study of voice onset time. **Journal of Phonetics**, [S. l.], v. 5, n. 4, p. 339-350, Oct. 1977.

BIALYSTOK, E. Explicit and implicit judgments of L2 grammaticality. **Language Learning**, [S. l.], v. 29, n. 1, p. 81–103, June 1979.

BOERSMA, P.; WEENINK, D. **Praat**. Version 6.3.10. Amsterdam, 2023. In: <https://www.praat.org>.

BROWNE, C.; CULLIGAN, B.; PHILLIPS, J. **The New General Service List**. 2013 Retrieved from <https://ngslprofiler.com/>. Accessed on 23 Oct. 2024.

CELCE-MURCIA, M.; BRINTON, D. M.; GOODWIN, J. M. **Teaching pronunciation: a course book and reference guide**. Cambridge: Cambridge University Press, 2010.

CHO, T.; LADEFOGED, P. Variations and universals in VOT. **Journal of phonetics**, [S. l.], v. 27, n. 2, p. 207-229, Apr. 1999.

CHODROFF, E.; GODFREY, J.; KHUDANPUR, S.; WILSON, C. Structured variability in acoustic realization: A corpus study of voice onset time in American English stops. In The Scottish Consortium for ICPhS (Ed.). Proceedings of the 18th international congress of phonetic sciences. **Proceedings [...]**. Glasgow, UK: the University of Glasgow. 2015.

CHODROFF, E.; WILSON, C. Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. **Journal of Phonetics**, [S. l.], v. 61, p. 30-47, Mar. 2017.

CLOPPER, C. G.; PISONI, D. B.; JONG, K. Acoustic characteristics of the vowel systems of six regional varieties of American English. **The Journal of The Acoustical Society of America**, [S. l.], v. 118, n. 3, p. 1661-1676, Sept. 2005.

COHEN, G. V. The VOT dimension: **A bi-directional experiment with English and Brazilian-Portuguese stops**. 96f. Dissertação (Mestrado em Língua Inglesa) – Universidade Federal de Santa Catarina, Florianópolis, 2004.

COHEN, J. A power primer. **Psychological Bulletin**, [S. l.], v. 112, n. 1, p. 155–159, July 1992. DOI: 10.1037//0033-2909.112.1.155

COHEN, J.; COHEN, P.; WEST, S. G.; AIKEN, L. S. **Applied multiple regression/correlation analysis for the behavioral sciences**. Mahwah, NJ: Erlbaum, 2003.

CUNNINGHAM, A.; CARROLL, J. (2015). Early predictors of phonological and morphological awareness and the link with reading: Evidence from children with different patterns of early deficit. **Applied Psycholinguistics**, [S. l.], v. 36, n. 3, p. 509-531, May 2015.

DEKEYSER, R. Implicit and explicit learning. In: DOUGHTY, C. J.; LONG, M. (Eds.). **The handbook of second language acquisition**. Oxford: Blackwell, 2003, p. 313–348.

DEKEYSER, R. **Practice in a second language**: Perspectives from applied linguistics and cognitive psychology. 1. ed. New York: Cambridge University Press, 2007.

DEKEYSER, R. Knowledge and skill in ISLA. *In*: LOEWEN, S.; SATO, M. (Eds.). **The Routledge handbook of instructed second language acquisition**. New York: Routledge, 2017. p. 15–32.

DEXHEIMER, B.; PRZYBYLA, A.; MURPHY, T.; AKPINAR, S.; SAINBURG, R. Reaction time asymmetries provide insight into mechanisms underlying dominant and non-dominant hand selection. **Experimental brain research**, [S. l.], v. 240, n. 10, p. 2791-802, Oct. 2022.

EHRI, L. C.; O'LEARY, R. Ensinar letras e consciência fonêmica como habilidades fundamentais ajuda as crianças de 4 e 5 anos a avançarem na leitura. *In*: SARGIANI, R. (Org.). **Alfabetização baseada em evidências**: Da ciência à sala de aula. Porto Alegre: Penso, 2022. p. 61-76.

ELLIS, N. The weak interface, consciousness, and form-focused instruction: Mind the doors. *In*: FOTOS, S.; NASSAJ, H. (Eds.). **Form-focused instruction and teacher education**: Studies in honor of Rod Ellis. Oxford: OUP, 2006. p. 17-34.

ELLIS, N. Implicit and explicit knowledge about language. *In*: CENOZ, J; HORNBERGER, N. H. (Eds.). **Encyclopedia of language and education**, v. 6: Knowledge about Language. 2. ed. New York: Springer, p. 1-13, 2008.

ELLIS, R. **The study of second language acquisition**. Oxford: Oxford University Press, 1994.

ELLIS, R. The Definition and Measurement of L2 Explicit Knowledge. **Language Learning**, [S. l.], v. 54, n. 2, p. 227–275, June 2004.

ELLIS, R. Measuring implicit and explicit knowledge of a second language: A psychometric study. **Studies in Second Language Acquisition**, [S. l.], v. 27, p. 141-172, June 2005.

ELLIS, R. Implicit and explicit learning, knowledge and instruction. *In*: ELLIS, R.; LOEWEN S.; ELDER C.; ERLAM R.; PHILP J.; REINDERS H. (Eds.), **Implicit and explicit knowledge in second language learning, testing and teaching**. Bristol: Multilingual Matters, 2009, p. 3-26.

FALKIEVICZ, D. JR. **The production of alveolar stops by multilingual speakers of Portuguese-Polish-English**. Thesis (Master's). Florianópolis: Universidade Federal de Santa Catarina, 2024.

FLEGE, J. E. Give Input a Chance! *In*: PISKE, Thorsten; YOUNG-SCHOLTEN, Martha (ed.). Input matters in SLA. Bristol: **Multilingual Matters**, 2009. p. 175-190.

FLEGE, J. E.; BOHN, O. The revised speech learning model (SLM-r). *In*: WAYLAND, Ratree (ed.). **Second Language Speech Learning**: Theoretical and empirical progress, Cambridge: Cambridge University Press, 2021, p. 3-83.

FLEGE, J. E.; FRIEDA, E. M.; NOZAWA, T. Amount of native-language (L1) use affects the pronunciation of an L2. **Journal of Phonetics**, [S. l.], p. 169-186, Apr. 1997.

FLEGE, J. E. MACKAY, L. R. A. PERCEIVING VOWELS IN A SECOND LANGUAGE. **Studies in Second Language Acquisition**, [S. l.], v. 26, n. 1, p. 1–34, Mar. 2004.

FLEGE, J. E.; MUNRO, M. J.; MACKAY, L. R. A. Factors affecting strength of perceived foreign accent in a second language. **The Journal of the Acoustical Society of America**, [S. l.], v. 97, n. 5, p. 3125-3134, May 1995.

FRANÇA, K. **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.

GEUDENS, A. Phonological Awareness and Learning to Read a First Language: Controversies and New Perspectives. **LESLLA Symposium Proceedings**, [S. l.], v. 1, n. 1, p. 25–43, 2006. DOI: 10.5281/zenodo.7980171. Available at: <https://lesllasp.journals.publicknowledgeproject.org/index.php/lesllasp/article/view/5533>. Accessed on 25 May 2025.

JASP TEAM. **JASP**. Version 0.19. Amsterdam, 2025. Retrieved from: <https://jasp-stats.org/>

JENKINS, J. J.; STRANGE, W.; POLKA, L. Not everyone can tell a “rock” from a “lock”: Assessing individual differences in speech perception. In: LUBINSKI, D. J.; DAWIS, R. V. (Eds.). **Assessing individual differences in human behavior**: New concepts, methods, and findings. Palo Alto, CA: Davies-Black Publishing, 1995. p. 297–325.

KENNEDY, S. Exploring the relationship between language awareness and second language use. **TESOL Quarterly**, [S. l.], v. 46, n. 2, p. 398-408, June 2012.

KENNEDY, S.; TROFIMOVICH, P. Language awareness and second language pronunciation: A classroom study. **Language Awareness**, [S. l.], v. 19, n. 3, p. 171-185. Dec. 2010.

KENT, R. D.; READ, C. **The Acoustic Analysis of Speech**. San Diego: Singular, 1992.

KIM, M. **Exploring the interface of explicit and implicit second-language knowledge: a longitudinal perspective**. 2020. Dissertation (Doctoral) – Boston University, Boston, 2020.

KIM, K., GODFROID, A. The interface of explicit and implicit second-language knowledge: A longitudinal study. **Bilingualism: Language and Cognition**, v. 26, p. 709–723, Jan. 2023.

KIVISTÖ-DE SOUZA, H. **Phonological awareness and pronunciation in a second language**. 2015. 500 f. Dissertation (Doctoral) - Doctoral Program in Applied Linguistics, Department of English and German Studies, University of Barcelona, Barcelona, 2015.

KIVISTÖ-DE SOUZA, H. Brazilian EFL Learners’ Awareness about L2 Phones: Is ‘MALL’ pronounced as ‘MAL’? **Trabalhos em Linguística Aplicada**, [S. l.], v. 56, n. 1, p. 235–258, Apr. 2017.

KIVISTÖ-DE SOUZA, H. Consciência Fonológica. *In*: KUPSKE, Felipe Flores; ALVES, Ubiratã Kickhöfel; LIMA JUNIOR, Ronaldo (Org.). **Investigando os sons de línguas não nativas**: Uma introdução. Campinas: Editora da ABRALIN, 2021. p. 153-173.

KIVISTÖ-DE SOUZA, H.; LINTUNEN, P. Thinking about your pronunciation: Examining phonological self-awareness with a novel task. *In*: HENDERSON, A.; KIRKOVA-NASKOVA, A. (Eds.), **Proceedings of the 7th International Conference on English Pronunciation**: Issues and Practices, Université Grenoble-Alpes, 2023. p. 138-148. <https://doi.org/10.5281/zenodo.8225603>

KLATT, D. Voice onset time, frication, and aspiration in word-initial consonant clusters. **Journal of Speech and Hearing Research**, [S. l.], v. 18, n. 4, p. 686-706, Dec. 1975.

KLEIN, S. **Estudo do VOT no português brasileiro**. Dissertação (Mestrado em Letras). Florianópolis: Universidade Federal de Santa Catarina, 1999.

KRASHEN, S. D. **Second language acquisition and second language learning**. 1. ed. Oxford: Pergamon Press, 1981.

KRASHEN, S. D. **The input hypothesis**: Issues and implications. 1. ed. Addison-Wesley Longman Ltd, 1985.

KUPISCH T.; BARTON D.; HAILER K.; KLASCHIK E.; STANGEN I.; LEIN T.; VAN DE WEIJER J. Foreign accent in adult simultaneous bilinguals. **Heritage Language Journal**, [S. l.], v. 11. n. 2, p.123-150, Mar. 2014.

KUPSKE, F. F. **Imigração, Atrito e Complexidade**: A Produção das Oclusivas Surdas Iniciais do Inglês e do Português por Sul-Brasileiros Residentes em Londres. Tese (Doutorado em Letras) Programa de Pós-Graduação em Letras, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2016.

KUPSKE, F. F.; LIMA JR., R. M. Integração a contextos de L2 dominantes e adaptabilidade fonológica de L1: Uma análise da produção das plosivas surdas do português brasileiro. **Organon**, Porto Alegre, v. 37, n. 73, p. 173-98, July 2022.

KUPSKE, F. F.; LIMA JR, R. M; ALVES U. K. A comparison between Full Time Equivalent and Length of Residence as measurements of time in bilingual speech research: the case of first language attrition. **Ilha do Desterro**, Florianópolis, v. 76, n. 3, p. 177-98, Sept. 2023.

KUPSKE, F. F.; OLIVEIRA, M. O desenvolvimento do padrão de Voice Onset Time das oclusivas surdas iniciais do inglês por aprendizes soteropolitanos: efeitos da instrução explícita. **Ilha do Desterro**, Florianópolis, v. 73, n. 3, p. 185-204, Nov. 2020.

KUPSKE, F. F.; PEROZZO, R. V. Efeitos multidirecionais da instrução explícita de pronúncia: O desenvolvimento de VOT longo em inglês (L2) e sua influência nas oclusivas do português brasileiro (L1). **Ilha do Desterro**, Florianópolis, v. 77, p. 1-23, Nov. 2024.

LADEFOGED, P.; JOHNSON, K. **A Course in Phonetics**. 6 ed. Boston: MA: Wadsworth/Cengage Learning, 2011.

LARSON-HALL, J. **A guide to doing statistics in second language research using SPSS and R**. New York: Routledge, 2015.

LEMHÖFER, K.; BROERSMA, M. Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. **Behavior Research Methods**. Nijmegen, v. 44, p. 325-343, Sept. 2012.

LIMA JR, R. M. **A influência da idade na aquisição da fonologia do inglês por brasileiros**. 2012. Tese (Doutorado) – Universidade de Brasília, Programa de Pós-Graduação em Linguística, Brasília (DF), 2012.

LIMA JR, R. M. A influência da idade na aquisição de seis vogais do inglês por alunos brasileiros. **Organon**, Porto Alegre, v. 30, n. 58, p. 15-31. July 2015. Universidade Federal do Rio Grande do Sul.

LIMA JR, R. M.; GARCIA, G. D. Diferentes análises estatísticas podem levar a conclusões categoricamente distintas. **Revista da ABRALIN**, [S. l.], v. 20, n. 1, p. 1–19, 2021. DOI: 10.25189/rabralin.v20i1.1790. Disponível em: <https://revista.abralin.org/index.php/abralin/article/view/1790>. Acesso em: 27 nov. 2025.

LISKER, L.; ABRAMSON, A. S. A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. **WORD**, [S. l.], v. 20, n. 3, p. 384–422, Jan. 1964.

LISKER, L.; ABRAMSON, A. S. Some effects of context on voice onset time in English stops. **Language and speech**, [S. l.], v. 10, n. 1, p. 1-28, Jan. 1967.

MENDES, C. F. **The perception of the English –s morpheme by Brazilian EFL learners**. 2017. Thesis (Mestrado em Inglês) – Programa de Pós-Graduação em Inglês, Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2017.

MILLER, J.; VOLAITIS, L. Effect of speaking rate on the perceptual structure of a phonetic category. **Perception & Psychophysics**, [S. l.], v. 46, p. 505–512, Nov. 1989.

MOYER, A. Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. **Studies in second language acquisition**, [S. l.], v. 21, n. 3, p 81-108, Mar. 1999.

MOYER, A. Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. In: PISKE, Thorsten; YOUNG-SCHOLTEN, Martha (Eds.). **Input matters in SLA**. Bristol: Multilingual Matters, 2009. p. 159-174.

MURPHY, K. R.; MYORS, B. **Statistical power analysis: A simple and general model for traditional and modern hypothesis tests**. 5. ed. Abingdon/New York: Routledge, 2023.

OAKHILL, J.; KYLE, F. The Relation between Phonological Awareness and Working Memory. **Journal of Experimental Child Psychology**, [S. l.], v. 75, n. 2, p. 152-164, Feb. 2000.

OTTO, S. E. From past to present: A hundred years of technology for L2 learning. *In*: CHAPPELLE, C. A.; SAURO, S. (Org.). **The Handbook of Technology and Second Language Teaching and Learning**. Hoboken: John Wiley & Sons, Inc., 2017. p. 10-25.

PISKE, T.; MACKAY, I.; FLEGE, J. Factors affecting degree of foreign accent in an L2: A review. **Journal of Phonetics**, [S. l.], v. 29, p. 191-215, Apr. 2001.

POSIT TEAM. **RStudio**: Integrated Development Environment for R. Version 5.4.2. Posit Software, PBC, Boston, MA. 2025. Retrieved from: <https://cran.r-project.org/>

PRESTES, S. P. D. C. **Produção de consoantes oclusivas iniciais do inglês por falantes nativos de PB**. Dissertação (Mestrado em Letras) - Universidade Federal do Paraná, Curitiba, 2013.

RAMÍREZ VERDUGO, D. A Study of Intonation Awareness and Learning in Non-native Speakers of English. **Language Awareness**, [S. l.], v. 15, n. 3, p. 141-159, Jan. 2006.

ROEHR, K.; GÁNEM-GUTIÉRREZ, G. A. The status of metalinguistic knowledge in instructed adult L2 learning. **Language Awareness**, [S. l.], v. 18, n. 2, p. 165-181. Aug. 2009.

ROEHR-BRACKIN, K. Explicit and implicit knowledge and learning of an additional language: A research agenda. **Language Teaching**, [S. l.], v. 57, n. 1, p. 68-86. Jan. 2024.

RYALLS, J.; ZIPPRER, A.; BALDAUFF, P. A preliminary investigation of the effects of gender and race on voice onset time. **Journal of Speech, Language, and Hearing Research**, [S. l.], v. 40, n. 3, p. 642-645, June 1997.

SAITO, K.; HANZAWA, K. The role of input in second language oral ability development in foreign language classrooms: A longitudinal study. **Language Teaching Research**, [S. l.], v. 22, n. 4, p. 398-417, July 2018.

SANCIER, M. L.; FOWLER, C. A. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. **Journal of Phonetics**, [S. l.], v. 27, n. 4, p. 421-436, 1997.

SCHERTZ, J.; HAWTHORNE, K. The effect of sentential context on phonetic categorization is modulated by talker accent and exposure. **The Journal of the Acoustical Society of America**, [S. l.], v. 143, n. 3, p. 231-236, Mar. 2018.

SCHMIDT, R. The role of Consciousness in Second Language Learning. **Applied Linguistics**. Oxford, v. 11, n. 2, p. 129-158. June. 1990.

SCHMIDT, R. Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. *In* SCHMIDT, R. (ed.). **Attention and Awareness**. Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center. 1995. p. 1-63.

SCHMIDT, R. Attention, Awareness, and Individual Differences in Language Learning. *In*: Fourth CLaSIC 2010 Individual Characteristics and Subjective Variables in Language Learning, 4., 2010, Singapore. **Proceedings** [...]. Singapore: National University of Singapore Centre for Language Studies, 2010. p. 721-737.

SCHOONMAKER-GATES, E. Measuring foreign accent in Spanish: How much does VOT really matter. *In*: WILLIS, Erik W. *et al.*, (Eds.). *Selected proceedings of the 6th conference on laboratory approaches to Romance phonology*, 2015. **Proceedings [...]**. Somerville, MA, 2015a. p. 95-105.

SCHWARTZHAUPT, B.; ALVES, U. K.; FONTES, A. B. The role of L1 knowledge on L2 speech perception: Investigating how native speakers and Brazilian learners categorize different VOT patterns in English. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 23, n. 2, p. 311, Sept. 2015.

SEARA, I. C.; NUNES, V. G.; LAZZAROTTO-VOLCÃO, C. **Fonética e Fonologia do Português**. 2. ed. Florianópolis: LLV/CCE/UFSC, 2011. 119 p.

SHOEMAKER, E. The exploitation of subphonemic acoustic detail in L2 speech segmentation. **Studies in Second Language Acquisition**, [*S. l.*], v. 36, n. 4, p. 709-731, Dec. 2014.

SUZUKI, Y.; DEKEYSER, R. The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. **Language Learning**, Michigan, v. 67, n. 4, p. 747-90, Dec. 2017.

TELÓ, C.; KIVISTÖ-DE SOUZA, H. The acquisition of a novel phonetic category in a foreign language setting: Input versus phonological awareness. **Ilha do Desterro**, Florianópolis. v. 76, n. 3, p. 709-731, Nov. 2023.

VANPATTEN, B. Attending to form and content in the input: An experiment in consciousness. **Studies in second language acquisition**, [*S. l.*], v. 12, n. 3, p. 287-301, Sept. 1990.

VENKATAGIRI, H. S.; LEVIS, J. M. Phonological awareness and speech comprehensibility: An exploratory study. **Language awareness**, [*S. l.*], v. 16, n. 4, p. 263-277, Nov. 2007.

WHITESIDE, S.; IRVING, C. Speakers' sex differences in voice onset time: A study of isolated word production. **Perceptual and motor skills**, [*S. l.*], v. 86, n. 2, p. 651-654, Apr. 1998.

WINN, M. B. Manipulation of voice onset time in speech stimuli: A tutorial and flexible Praat script. **The Journal of the Acoustical Society of America**, [*S. l.*], v. 47, n. 2, p. 852-866, Feb. 2020.

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

YAVAS, M. **Applied English Phonology**. 2. ed. Oxford: Blackwell, 2011.

ZIMMER, M. C.; SILVEIRA, R.; ALVES, U. K. **Pronunciation Instruction for Brazilians: Bringing Theory and Practice Together.** Newcastle upon Tyne. Cambridge: Cambridge Scholars Publishing (CSP), 2009.

**APPENDIX A – TOKENS INCLUDED IN THE FAMILIARIZATION SESSION OF
THE *ERROR DETECTION TEST***

Token	Accuracy	VOT duration (ms) or initial sound
cap	accurate	99
salt	inaccurate	[z]
meat	accurate	[m]
pet	accurate	89
cap	inaccurate	38.93
top	accurate	98
salt	accurate	[s]
jury	accurate	[dʒ]
far	accurate	[f]
pet	inaccurate	15.13

Source: the author.

The familiarization session followed this fixed order of tokens for all participants, with no repeated tokens in sequence and with three distractors ('meat', 'salt', and 'far') taken from the test measuring explicit awareness, apart from 'jury', which is also included in this task.

APPENDIX B – TEST TOKENS INCLUDED IN THE *ERROR DETECTION TEST*

Token	VOT duration (ms) of the target words			
	Female		Male	
	Accurate	Inaccurate	Accurate	Inaccurate
pig	89	15.13	89	15.13
peace	89	15.13	89	15.13
pool	89	15.13	89	15.13
pack	89	15.13	89	15.13
pause	89	15.13	89	15.13
tell	98	17.88	98	17.88
ten	98	17.88	98	17.88
tool	98	17.88	98	17.88
task	98	17.88	98	17.88
tall	98	17.88	98	17.88
kid	99	38.93	99	38.93
keep	99	38.93	99	38.93
cool	99	38.93	99	38.93
cat	99	38.93	99	38.93
call	99	38.93	99	38.93
	Initial sound of the distractors			
job	[dʒ]	[ʒ]	[dʒ]	[ʒ]
joke	[dʒ]	[ʒ]	[dʒ]	[ʒ]
jump	[dʒ]	[ʒ]	[dʒ]	[ʒ]
jail	[dʒ]	[ʒ]	[dʒ]	[ʒ]
joy	[dʒ]	[ʒ]	[dʒ]	[ʒ]
June	[dʒ]	[ʒ]	[dʒ]	[ʒ]
judge	[dʒ]	[ʒ]	[dʒ]	[ʒ]
joint	[dʒ]	[ʒ]	[dʒ]	[ʒ]
gene	[dʒ]	[ʒ]	[dʒ]	[ʒ]
gym	[dʒ]	[ʒ]	[dʒ]	[ʒ]
Total number	25	25	25	25

Source: the author.

APPENDIX C – BACKGROUND QUESTIONNAIRE

Questionário 1:

Prezado(a) participante,

Este questionário tem como objetivo obter informações a serem utilizadas na análise dos dados deste estudo. Nem sua identidade nem qualquer informação que possa revelá-la serão divulgadas.

1. Nome completo:

2. Idade:

3. Gênero:

- ☐ Feminino
- ☐ Masculino
- ☐ Outro
- ☐ Prefiro não dizer

4. Local de nascimento (cidade/estado):

5. Onde você reside atualmente (cidade/estado):

5.1. Há quanto tempo?

6. Ocupação:

6.1. Se for estudante, o que está estudando?

7. Você estuda ou já estudou Letras?

- ☐ Sim
- ☐ Não

7.1. O que (marque todas que se aplicam)?

- ☐ Graduação em Letras Inglês em andamento.
- ☐ Graduação em Letras (outro idioma) em andamento.
- ☐ Mestrado em Letras/Linguística em andamento.
- ☐ Doutorado em Letras/Linguística em andamento.
- ☐ Graduação em Letras Inglês concluída.
- ☐ Graduação em Letras (outro idioma) concluída.
- ☐ Mestrado em Letras/Linguística concluído.
- ☐ Doutorado em Letras/Linguística concluído.

8. Você possui alguma deficiência de fala ou audição?

- ☐ Sim
- ☐ Não

8.1. Qual(is)?

9. Qual é a sua mão dominante?

- ☐ Direita (Destro/a)
- ☐ Esquerda (Canhoto/a)

10. Você já foi para o exterior?

- ☐ Sim
- ☐ Não

10.1. Onde e por quanto tempo (cada)?

11. Com que idade você começou a estudar/aprender inglês?

12. De que forma você aprendeu inglês, principalmente?

- ☐ Escola
- ☐ Cursinho de idiomas
- ☐ Mídia (ex.: filmes, música, TV, jogos)
- ☐ Intercâmbio
- ☐ Amigos / Socialização
- ☐ Outros.

12.1. Como?

13. Você já fez aulas de inglês (Ex.: escola, cursinho)?

- ☐ Sim
- ☐ Não

13.1. Por quanto tempo no total (somando todas as vezes)?

13.2. Como eram as aulas de inglês que você teve?

- ☐ Ensino focado principalmente em gramática (ou seja, muito tempo era dedicado ao estudo da gramática).
- ☐ Ensino focado principalmente na comunicação (ou seja, a maior parte do tempo era dedicada à comunicação em inglês).
- ☐ Uma mistura de ensino voltado para gramática e para a comunicação.

13.3. Com que frequência a pronúncia era abordada nas suas aulas de inglês?

- ☐ Toda aula
- ☐ Quase toda aula
- ☐ Metade das aulas
- ☐ Poucas aulas
- ☐ Nenhuma aula

14. Em **UMA SEMANA** típica, quantas **horas** no total você faz cada uma das seguintes atividades abaixo **EM INGLÊS**:

14.1. Assisto a filmes e/ou séries sem legendas. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.2. Assisto a filmes e/ou séries com legendas. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.3. Escuto podcasts. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.4. Assisto a vídeos (ex. TikTok, Reels, YouTube). **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.5. Assisto a lives no Instagram, Twitch, YouTube, etc. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.6. Participo de aulas presenciais e online com professores que falam a maior parte do tempo em inglês. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.7. Participo de reuniões, palestras ou eventos. **(em inglês)**.

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.8. Escuto áudios de livros didáticos e outros materiais. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.9. Converso (falo) em inglês com outros falantes não nativos de inglês. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.10. Converso (falo) em inglês com falantes nativos de inglês. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.11. Gravo áudios e/ou vídeos (ex. WhatsApp, TikTok). **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.12. Escrevo publicações e comentários em redes sociais. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.13. Troco mensagens de texto (ex. WhatsApp, e-mail). **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.14. Escrevo diários, fanfics e/ou livros. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.15. Escrevo atividades de produção escrita (ex. redação, homework). **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.16. Leio publicações e comentários em redes sociais. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.17. Leio livros, fanfics e/ou jornais. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

14.18. Leio artigos acadêmicos. **(em inglês)**

0 horas	Menos de 1 hora	1 - 2 horas	3 - 5 horas	6 - 10 horas	Mais de 10 horas
---------	-----------------	-------------	-------------	--------------	------------------

15. Pense nas 5 pessoas com as quais você mais mantém contato no dia a dia. Indique o idioma que vocês usam para se comunicar:

	Somente português	Somente inglês	Metade português Metade inglês	Mais português	Mais inglês	Outro
Pessoa 1	☐	☐	☐	☐	☐	☐
Pessoa 2	☐	☐	☐	☐	☐	☐
Pessoa 3	☐	☐	☐	☐	☐	☐
Pessoa 4	☐	☐	☐	☐	☐	☐
Pessoa 5	☐	☐	☐	☐	☐	☐

16. De iniciante (nº01) a avançado (nº10), qual nível você considera que possui em inglês?

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

17. Você já realizou algum teste de proficiência? (ex.: FCE, TOEFL, TOEIC, Duolingo Test)?

- ☐ Sim
☐ Não

17.1. Qual foi o teste e qual a sua pontuação/nível?

18. Estime seu nível de proficiência em inglês nas seguintes habilidades:

a. Compreensão oral (listening):

Básico	Intermediário	Intermediário avançado	Avançado	Nativo
--------	---------------	---------------------------	----------	--------

b. Fala (speaking):

Básico	Intermediário	Intermediário avançado	Avançado	Nativo
--------	---------------	---------------------------	----------	--------

c. Leitura (reading):

Básico	Intermediário	Intermediário avançado	Avançado	Nativo
--------	---------------	---------------------------	----------	--------

d. Escrita (writing):

Básico	Intermediário	Intermediário avançado	Avançado	Nativo
--------	---------------	---------------------------	----------	--------

19. Você fala algum outro idioma estrangeiro?

☐ Sim

☐ Não

19.1. Quais idiomas e qual seu nível de proficiência?

20. Você já participou de algum curso específico de pronúncia em inglês?

☐ Sim

☐ Não

20.1. Qual era o nome do curso, onde foi feito e qual a duração?

21. Caso haja alguma outra informação que considere relevante para esta pesquisa em relação ao seu uso de inglês, por favor, comente abaixo.

Source: the author.

**APPENDIX D – INTERNAL CONSISTENCY OF THE QUESTIONNAIRE
MEASURING EXPLICIT AWARENESS**

Section	Item type		Number of items	Average inter-item correlation	95% CI	
					Lower	Upper
Audio- based	Accurate	BP	3	-	-	-
		English	3	-0.036	-0.086	-0.028
	Inaccurate	BP	6	-	-	-
		English	6	0.152	0.075	0.233
Text- based	First set of statements		3	0.06	-0.149	0.297
	Second set of statements		3	0.206	0.006	0.421

Source: the author.

Key: BP = Brazilian Portuguese; CI = confidence interval.

Note: the analysis for the items in BP from the *Audio-based* section could not be run due to some items that had a variance equal to zero.

APPENDIX E – FREQUENCY AND RELATIVE FREQUENCY OF THE CERTAINTY-RATINGS

Item type	Response accuracy	Certainty-ratings (%)			All
		Certain	Semi-certain	Uncertain	
BP accurate	Correct	106 (95.5)	3 (2.7)	0 (0)	111 (100)
	Incorrect	1 (0.9)	0 (0)	1 (0.9)	
	All	107 (96.4)	3 (2.7)	1 (0.9)	
BP inaccurate	Correct	28 (12.61)	22 (9.91)	4 (1.8)	222 (100)
	Incorrect	158 (71.17)	8 (3.6)	2 (0.9)	
	All	186 (83.78)	30 (13.51)	6 (2.7)	
English accurate	Correct	90 (81.08)	15 (13.51)	2 (1.8)	111 (100)
	Incorrect	2 (1.8)	2 (1.8)	0 (0)	
	All	92 (82.88)	17 (15.32)	2 (1.8)	
English inaccurate	Correct	30 (13.51)	25 (11.26)	3 (1.35)	222 (100)
	Incorrect	120 (54.05)	41 (18.47)	3 (1.35)	
	All	150 (67.57)	66 (29.73)	6 (2.7)	

Source: the author.

Key: BP = Brazilian Portuguese.

Calculation example: English accurate – Correct – Certain

Frequency = 90

Relative frequency = $(90 / 111) * 100 = 81.08$

APPENDIX F – DATA ANALYSIS OF THE *L2 INPUT SCORE*

Statement		Language skill
Q14.1	I watch movies and/or series without subtitles.	Listening
Q14.2	I watch movies and/or series with subtitles.	Reading
Q14.3	I listen to podcasts.	Listening
Q14.4	I watch videos (e.g., TikTok, Reels, YouTube).	Listening
Q14.5	I watch live streams on Instagram, Twitch, YouTube, etc.	Listening
Q14.6	I take classes with teachers who speak mostly in English.	Listening
Q14.7	I participate in meetings, lectures, or events.	Listening
Q14.8	I listen to audio from textbooks and other materials.	Listening
Q14.9	I talk (speak) in English with other non-native English speakers.	Speaking
Q14.10	I talk (speak) in English with native English speakers.	Speaking
Q14.11	I record audio and/or video (e.g., WhatsApp, TikTok).	Speaking
Q14.12	I write posts and comments on social media.	Writing
Q14.13	I exchange text messages (e.g., WhatsApp, email).	Writing
Q14.14	I write diaries, fanfiction, and/or books.	Writing
Q14.15	I write assignments (e.g., essays, homework).	Writing
Q14.16	I read posts and comments on social media.	Reading
Q14.17	I read books, fan fiction, and/or newspapers.	Reading
Q14.18	I read academic articles.	Reading
Internal consistency		
n	Statements included	Cronbach's Alpha 95% CI
11	Oral skills (listening + speaking)	.49 [.22, .68]
7	Oral skills (except statements Q14.6, Q14.7, Q14.8, and Q14.11 that correlated negatively with the scale)	.61 [.37, .78]
7	Written skills (reading + writing)	.65 [.47, .77]
6	Written skills (except statement Q14.13 that correlated negatively with the scale)	.70 [.55, .81]
18	All statements	.71 [.57, .82]
14	All statements (except statements Q14.6, Q14.7, Q14.8, and Q14.11 that correlated negatively with the scale)	.76 [.63, .85]
13	All statements (except statements Q14.6, Q14.7, Q14.8, and Q14.11 that correlated negatively with the scale, and statement Q14.14 that showed the weakest correlation with the overall scale)	.77 [.64, .86]

Source: the author.

Key: n = number of items; CI = confidence interval.

Note: the statements were answered on a 6-point Likert-scale (0 hours – Less than 1 hour – 1-2 hours – 3-5 hours – 6-10 hours – More than 10 hours).

APPENDIX G – INDIVIDUAL SCORES

Participant ID	<i>Implicit Awareness Score</i>			<i>Explicit Awareness Score</i>		
	Accurate tokens	Inaccurate tokens	All	Text-based section	Audio-based section	All
p01	93,33	26,67	60,00	33,33	50,00	41,67
p02	100,00	6,67	53,33	83,33	33,33	58,33
p03	83,33	33,33	58,33	16,67	33,33	25,00
p04	96,67	6,67	51,67	16,67	33,33	25,00
p05	93,33	10,00	51,67	0,00	0,00	0,00
p06	100,00	0,00	50,00	16,67	50,00	33,33
p07	100,00	0,00	50,00	83,33	33,33	58,33
p08	100,00	0,00	50,00	16,67	0,00	8,33
p09	100,00	0,00	50,00	33,33	16,67	25,00
p10	93,33	13,33	53,33	16,67	16,67	16,67
p11	100,00	0,00	50,00	50,00	0,00	25,00
p12	96,67	20,00	58,33	50,00	50,00	50,00
p13	96,67	26,67	61,67	66,67	33,33	50,00
p14	90,00	10,00	50,00	16,67	16,67	16,67
p15	70,00	23,33	46,67	50,00	16,67	33,33
p16	96,67	26,67	61,67	33,33	16,67	25,00
p17	90,00	16,67	53,33	66,67	50,00	58,33
p18	76,67	23,33	50,00	50,00	0,00	25,00
p19	93,33	10,00	51,67	50,00	50,00	50,00
p20	100,00	3,33	51,67	33,33	0,00	16,67
p21	96,67	20,00	58,33	16,67	66,67	41,67
p22	96,67	6,67	51,67	33,33	33,33	33,33
p23	86,67	13,33	50,00	66,67	16,67	41,67
p24	93,33	46,67	70,00	0,00	66,67	33,33
p25	96,67	10,00	53,33	16,67	33,33	25,00
p26	36,67	50,00	43,33	50,00	0,00	25,00
p27	96,67	0,00	48,33	33,33	0,00	16,67
p28	100,00	16,67	58,33	33,33	33,33	33,33
p29	96,67	40,00	68,33	33,33	66,67	50,00
p30	96,67	6,67	51,67	33,33	0,00	16,67
p31	100,00	0,00	50,00	16,67	0,00	8,33
p32	100,00	0,00	50,00	50,00	33,33	41,67
p33	96,67	3,33	50,00	33,33	16,67	25,00
p34	96,67	0,00	48,33	33,33	0,00	16,67
p35	100,00	10,00	55,00	16,67	0,00	8,33

p36	66,67	30,00	48,33	16,67	50,00	33,33
p37	73,33	23,33	48,33	33,33	50,00	41,67

Source: the author.

Note: The *Implicit Awareness Score* is the percentage of correct responses to the accurate (column 2) and inaccurate tokens (column 3) of the *Error Detection Test*, and the *Explicit Awareness Score* is the percentage of correct responses to the 3 statements of the first and the second set of statements referring to VOT from the *Text-based section* (column 5), plus the 6 audio items in English displaying inaccurate VOT values from the *Audio-based section* (column 6), comprising 12 items in total (column 7). All numbers are percentages of correct responses.