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

Pietra Cassol Rigatti

Haitian Creole speakers reading and listening to Brazilian Portuguese words: Is there a cognate facilitation effect?

Florianópolis
2024

Pietra Cassol Rigatti

Haitian Creole speakers reading and listening to Brazilian Portuguese words: Is there a cognate facilitation effect?

Tese submetida ao Programa de Pós-Graduação em Linguística da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Doutora em Linguística, na área de concentração Aquisição, aprendizagem e processamento da linguagem.

Orientadora: Prof.ª Dr.ª Mailce Borges Mota
Co-orientador: Prof. Dr. Kenneth Pugh

Florianópolis

2024

Rigatti, Pietra Cassol
Haitian Creole speakers reading and listening to Brazilian Portuguese words : Is there a cognate facilitation effect? / Pietra Cassol Rigatti ; orientadora, Mailce Borges Mota, coorientador, Kenneth Richard Pugh, 2024.
339 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Programa de Pós-Graduação em Linguística, Florianópolis, 2024.

Inclui referências.

1. Linguística. 2. crioulo haitiano. 3. efeito cognato. 4. acesso lexical. 5. língua minoritária. I. Mota, Mailce Borges . II. Pugh, Kenneth Richard. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Linguística. IV. Título.

Pietra Cassol Rigatti

Haitian Creole speakers reading and listening to Brazilian Portuguese words: Is there a cognate facilitation effect?

O presente trabalho em nível de Doutorado foi avaliado e aprovado, em 19 de abril de 2024, pela banca examinadora composta pelos seguintes membros:

____

Prof.ª Dra. Ana Beatriz Arêas da Luz Fontes
Universidade Federal do Rio Grande do Sul

____

Prof. Dr. Augusto Buchweitz
University of Connecticut

____

Prof.ª Dra. Roberta Pires de Oliveira
Universidade Federal de Santa Catarina

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

GOV.BR
Documento assinado digitalmente
HERONIDES MAURILIO DE MELO MOURA
Data: 17/07/2024 14:22:46-0300
CPF: ***.870.274-**
Verifique as assinaturas em https://v.ufsc.br

____

Coordenação do Programa de Pós-Graduação

ICP-Edu
Documento assinado digitalmente
Mailce Borges Mota
Data: 17/07/2024 14:13:17-0300
CPF: ***.818.537-**
Verifique as assinaturas em https://v.ufsc.br

____

Prof.ª Dra. Mailce Borges Mota.
Orientadora

Florianópolis, 2024.

ACKNOWLEDGEMENTS

I would like to thank Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for supporting this research study with funding in two periods of my PhD: funding from CAPES PROEX (process number 88887.374880/2019-00), from August 2019 to August 2023. This funding was paused from March 2020 to December 2020 while I was receiving funding from CAPES Programa Geral de Cooperação Internacional (process number 88887.474589/2020-00) from March 2020 to January 2021. During those 11 months, I was a visiting scholar at Haskins Laboratories at Yale University, doing what we call in Portuguese *doutorado sanduíche*, when I had many opportunities to learn from and to collaborate with this cooperation program. I would also like to thank the Programa de Pós-Graduação em Linguística (PPGLin) and the Universidade Federal de Santa Catarina (UFSC) for promoting and believing in my education and career.

I would like to thank Prof. Dr. Mailce Borges Mota for sharing her knowledge and experience with me, since 2019, and with everyone who has joined the Laboratório da Linguagem e Processos Cognitivos (LabLing). Prof. Mailce offered me many opportunities for academic and personal growth, guided me throughout this journey, advised me whenever I needed it, and allocated funding from her PQ/CNPQ grant to the data collection of this study. I am forever grateful for her teachings, attention, and time.

I would like to thank Prof. Dr. Kenneth Pugh for welcoming me to Haskins Labs in March 2020 and for making sure I could interact and learn with the amazing group of researchers despite the appalling pandemic the world was suffering. I also thank Prof. Dr. Augusto Buchweitz for the partnership between Haskins Labs and the CAPES PGCI project. Those 11 months were precious for me as a researcher and as a person, and I am forever grateful for them. The PGCI project was coordinated by Prof. Dr. Augusto Buchweitz from 2017 to 2021 and had the LabLing as one of its research collaborators. This collaboration allowed my visit to the Haskins Laboratories and this thesis is, in great part, a result of the long-standing collaboration between Dr. Buchweitz and Dr. Mota. Projects such as the PGCI are invaluable for advancing science and for improving researcher's skills.

I would like to thank everyone who I have met at LabLing since 2019. You all welcomed me to a new city and inspired me to keep moving forward. I learned so much from you all and I am lucky to call some of you friends now. Conducting research with your help is more efficient than alone.

I would like to thank Prof. Dr. Nelson Canzian da Silva, from the Physics Department (FSC) at UFSC. Prof. Nelson kindly donated his knowledge, skills, and time to help me and my colleagues at LabLing with programming our online experiments on web browsers using Java Script and the jsPsych library. I am extremely grateful for it. Also, I was aided in many moments by João Coelho, who patiently helped me solve issues during data analyses and contributed to make my coding strategies more efficient. When it comes to data analyses, I would also like to thank Noam Siegelman (Hebrew University of Jerusalem), who I met at Haskins Laboratories and who kindly offered his time to help me with analyses using mixed-effect linear models. I am indebted to PGCI, Dr. Buchweitz, Dr. Mota and Dr. Pugh for allowing me to learn from such brilliant researchers.

I would like to thank the Florianópolis Department of Education and the State of Santa Catarina Department of Education in the Great Florianópolis region for collaborating and contributing to this dissertation study. I also thank the administration staff at each school I visited for their time and efforts during data collection. Their help was key to conducting this study. And if science can be considered a public good, it is also of public interest.

I would like to thank every participant who offered their time and energy for completing the tasks in this dissertation. Mèsi anpil! I hope my work helps increasing the general interest in studying Haitian Creole in the future.

I would like to thank Murilo, my partner and husband, who supported me at all times. You were always there for me even when we were in different countries. I love you so much.

I would like to thank my parents, my family, and all my friends for always rooting for me and lifting my spirits up despite my insecurities. I would like to thank my therapist for helping me cope with these insecurities. And I would also like to thank Jane Snaider, who welcomed me and helped me so many times during 2020 in Connecticut.

Truly, this dissertation is a collective effort as the scientific endeavor should be.

RESUMO

O objetivo geral desta tese foi investigar os efeitos da interação entre as línguas crioulo haitiano (CH) e português brasileiro (PB), da consciência fonológica (CF) em PB e dos hábitos de leitura em PB no processamento de língua em PB como uma segunda língua por falantes nativos de CH. Para tanto, dois estudos foram realizados separadamente: o Estudo 1 teve como participantes crianças e adolescentes matriculados em escolas públicas e investigou o efeito cognato entre CH e PB e a influência da CF em PB, e o Estudo 2 teve como participantes adultos e investigou o efeito cognato entre CH e PB e a influência dos hábitos de leitura em PB. No Estudo 1, 48 participantes completaram 5 tarefas e 1 questionário: uma tarefa de decisão lexical (DL) visual em PB, uma tarefa de DL auditiva em PB, um teste de CF em PB, um teste de identificação de letras, um teste de vocabulário receptivo em CH e um questionário de histórico linguístico. As tarefas de DL continham 60 palavras cognatas e 60 palavras não cognatas, únicas por tarefa, e 120 pseudopalavras. As palavras foram apresentadas aleatoriamente. Nos dois Estudos, as análises estatísticas foram realizadas no R com modelos lineares de efeitos mistos. No Estudo 1, os modelos não mostraram qualquer efeito significativo das palavras cognatas, nem na tarefa visual nem na auditiva, nem nos tempos de resposta (TR) nem nos acertos. A idade e o tamanho de palavra não pareceram influenciar o resultado nulo. Houve efeito significativo de CF nos acertos na tarefa de DL visual, que foi discutido com evidências de uma relação mútua entre CF, aprender a ler e desenvolvimento de habilidades de leitura e escrita. No Estudo 2, 35 participantes completaram 4 tarefas e 2 questionários: uma tarefa de DL visual em PB, uma tarefa automonitorada de compreensão de frases (CFr) visual em PB, uma tarefa automonitorada de CFr auditiva em PB, um teste de vocabulário receptivo em CH, um questionário sobre hábitos de leitura em PB e um questionário de histórico linguístico. As tarefas de CFr continham 40 frases em que a quinta palavra era cognata e de 40 frases em que a quinta palavra era não cognata, únicas por tarefa. Após cada bloco fixo de 5 frases, uma pergunta de compreensão sobre a frase anterior era visualmente apresentada. Blocos foram apresentados aleatoriamente. No Estudo 2, os modelos estatísticos não mostraram qualquer efeito significativo das palavras cognatas no reconhecimento de palavras isoladas ou em frase, visual ou auditivamente, nos TRs ou nos acertos. Não houve efeito significativo dos hábitos de leitura em nenhum modelo. O tamanho das palavras não parece ter influenciado o resultado nulo. A ausência de efeito cognato nos dois estudos foi discutida considerando estudos sobre coativação de línguas. Uma explicação especulativa foi proposta. Ela cogita aspectos da Hipótese da Qualidade Lexical, da abordagem da Ativação entre Línguas e da Abordagem de Aprendizado, incluindo influências ambientais no desenvolvimento linguístico, como nível socioeconômico, e diversidade de insumo linguístico.

Palavras-chave: crioulo haitiano; português brasileiro; língua minoritária; acesso lexical; efeito cognato.

ABSTRACT

The general objective of this dissertation was to investigate the effects of cross-linguistic interaction of HC and BP, of BP phonological awareness and of reading habits in BP on language processing in BP as a second language by native speakers of HC. In order to do so, two separate studies were carried out: Study 1 had children and teenagers enrolled in public schools as participants and investigated the cognate effect across HC and BP and the influence of phonological awareness in BP, and Study 2 had adults as participants and investigated the cognate effect across HC and BP and the influence of reading habits in BP. In Study 1, 48 participants completed five tasks and one questionnaire: a visual lexical decision (LD) task in BP, an auditory LD task in BP, a BP phonological awareness test, a letter identification test in BP, a HC receptive vocabulary test, and a language history questionnaire. The LD tasks consisted of unique 60 cognate words across HC and BP, 60 noncognate, and 120 pseudowords. Word presentation was randomized. In both studies, statistical analyses were ran in R, and inferential analyses used mixed-effect linear models. In Study 1, models showed no significant effect of cognate words during either spoken or written word recognition on either RTs or accuracy. Further analyses indicated that age and word length do not seem to have influenced the null result. There was a significant effect of phonological awareness on accuracy rates in the written LD task, which was discussed in relation to evidence of a mutual relationship between phonological awareness, learning to read and developing reading and writing skills. In Study 2, 35 participants completed four tasks and two questionnaires: a visual LD task in BP, a visual self-paced sentence comprehension (SC) task in BP, a self-paced SC task in BP, a HC receptive vocabulary test, a reading habits in BP questionnaire, and a language history questionnaire. The SC tasks were composed of unique 40 sentences containing a cognate word across HC and BP as the fifth word and of 40 containing a noncognate word as the fifth word. After every fixed 5-sentence block, a comprehension question about the last sentence was visually presented. Block presentation was randomized. In Study 2, models showed no significant effect of cognate words during either spoken or written word recognition in isolation or embedded in sentences on either RTs or accuracy. There was no significant effect from reading habits in any of the other models. Further analyses indicated that word length does not seem to have influenced this null result. The absence of any cognate effect in both studies was discussed in relation to studies investigating cross-linguistic influences. A speculative explanation is proposed. It considers aspects from the Lexical Quality Hypothesis, the Cross-language Activation account, and the Learning Account, as well as environmental influences on language development, such as socioeconomic status, and diversity of linguistic input.

Keywords: Haitian Creole; Brazilian Portuguese; minority language; lexical access; cognate effect.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BP	Brazilian Portuguese
HC	Haitian Creole
L1	First language
L2	Second language
LD	Lexical decision
LHQ	Language History Questionnaire
PA	Phonological awareness
RH	Reading habits
RT	Response time
SC	Sentence comprehension
SES	Socioeconomic status
SLD	Spoken lexical decision
SSC	Spoken sentence comprehension
UFSC	Universidade Federal de Santa Catarina
WLD	Written lexical decision
WSC	Written sentence comprehension

1 INTRODUCTION

This dissertation investigates the cognate facilitation effect on visual and auditory word recognition of Brazilian Portuguese (BP) by native speakers of Haitian Creole (HC), a language pair not studied before in Psycholinguistics. The cognate facilitation effect is identified when bilingual or multilingual participants display faster response times (RT) and/or higher accuracy rates after being presented with cognate words in comparison with noncognate words. Most of the literature on the topic shows facilitatory effects. However, the impact of cognate words is complex and may lead to interference/inhibition during processing. Both the facilitation or the interference effect which may arise from cognate words indicate that the lexicons of bilingual or multilingual people are interconnected and simultaneously active at least to some extent. In addition, very few studies about the cognate effect dedicate themselves to test this phenomenon with more diverse bilingual experiences and languages. Thus, the general objective of this dissertation is to investigate the effects of cross-linguistic interaction between Haitian Creole (HC) and Brazilian Portuguese (BP) on lexical access during comprehension of spoken and written BP by native speakers of HC.

In the remainder of this introduction, I will preface the subject of bilingualism and I will present some linguistic phenomena common to people who speak more than one language. Then, I will address the lack of research on less studied language pairs and on varied bilingual experiences. Next, I will explain why investigating cross-language influences across HC and BP is relevant and I will comment on how critical language development and learning to read in a majority language are when your everyday life depends on it. At the end, I will list the research questions and objectives for this dissertation.

There are at least 6,359 languages in the world according to the statistics available in the LangScape website, which is an interactive online map created by the Maryland Language Science Center (University of Maryland, 2018). This map allows you to zoom in and out of different parts of the globe and discover what languages are spoken there. It illustrates the argument made by Grosjean (2013) that the ratio of languages and countries is not proportional: there are many more languages in the world than there are countries. He also mentioned (Grosjean, 2021) that in Indonesia there are speakers of 722 languages, of 445 in India, and of 207 in Australia. These figures make it easy to conclude that bilingualism and multilingualism are much more common than we might think.

Grosjean (2021) has estimated that at least half of the world’s population has knowledge of two languages or dialects and uses them. Surely, this estimate will vary according to one's definition of bilingualism. For example, one definition is that a bilingual person has "native-like control of two languages" (Bloomfield, 1933, p. 56). More recent conceptualizations posit that a bilingual is someone "who actively uses two languages" (Kroll *et al.*, 2015, p. 1), which does not necessarily imply native-likeness. On the contrary, today the bilingual experience is understood to be so diverse that researchers recommend not simply considering "bilingual" as a category anymore (Poarch; Bialystok, 2015).

Bilingualism has been studied at least since 1923, when Frank Smith published his paper "Bilingualism and Mental Development". In the same year, D. J. Saer published "The Effect of Bilingualism on Intelligence". The first attempts to review and organize the field of bilingualism were Weinreich's (1953) "Languages in Contact" and Haugen's (1956) "Bilingualism in the Americas: A Bibliography and Research Guide" (Mackey, 1958). These first studies focused on describing the influences of language contacts and the language skills of monolinguals and bilinguals. Then, the interest was extended to cognitive factors associated with bilingualism and to language teaching strategies. Only in the 1970's research became consolidated as studies started investigating how words are learned and stored in memory (Souza, 2020, p. 182).

During the first decades, the impression was that bilingualism only had negative outcomes. Smith (1923) and Saer (1923) observed that bilinguals had inferior performance in tasks in comparison with monolinguals, which led them to conclude that having knowledge of two languages could make people confused. Darcy (1953, p. 50) conducted a review of the literature about the effects of bilingualism on intelligence and observed that most studies showed that "bilinguists suffer from a language handicap when measured by verbal tests of intelligence". As Haugen (1949, p. 272) mentioned, "bilingualism was in disrepute" at the time. Indigenous languages were not even added to conferences about languages other than English until 1940 (Haugen, 1949, p. 276). The study by Lerea and Kohut (1961) might be among the first ones which showed some bilingual advantage in behavioral language tasks.

Since then, much has been understood about what bilingualism entails. For example, since Preston and Lambert (1969), researchers have been extensively observing that the two languages of a bilingual person are activated in parallel to some extent in their minds. This means that one language is able to have some influence over the other one and that the organization of the bilingual language system is somewhat different from the monolingual one. Another phenomenon which seems very common in bilingualism is language switching (Poplack, 1980). Bilinguals alternate languages, apparently with no difficulty, depending on the context of use and the people they are talking to. In general, this can also be called code-mixing or code-switching and it is not a random occurrence. This language alternation happens following linguistic patterns, which indicates that switching implies linguistic skill. A third example of characteristics of bilingualism is the transfer of language skills. This is not an automatic phenomenon, and not all bilinguals display it. However, it has been documented and seems to be an indirect relationship of individual differences across languages. Sparks (2012) presented a compilation of studies which suggest that L1 skills may be transferred or at least may have some impact on L2 skills.

These characteristics show that bilingualism and multilingualism come in all shapes and sizes and vary in many aspects. A person may acquire two mother tongues at the same time and be considered a simultaneous bilingual while another person may be exposed to the second language after the first one is consolidated and be considered a sequential bilingual. Also, there are societies in which a major language organizes the government and the educational system while a minority language is used at home and around friends. Differently, there are societies structured with a major language and an international one. There are also societies regulated by only one official language but influenced by international ones. These different scenarios imply varied contexts of language use for bilingual people. Further, the diverse circumstances in which bilingualism emerges may require specific levels of proficiency in certain language modalities. It is possible that a bilingual is only required to read in one of their languages or that they were only exposed to a writing system in one of them. Similarly, bilinguals' comprehension abilities may be great in both languages, but their production skills might be advanced in only one. In addition, there are societies in which two very dissimilar varieties of a linguistic system — one usually spoken and the other, written — occupy specific everyday functions: the diglossia (Fishman, 1967).

Despite all this diversity in types of bilingualism, most research has focused on few languages and populations. A quick search for "bilingualism" and the name of a language on the Web of Science Core Collection showed us these numbers: 11,785 entries for English, 2,744 for Spanish, 1,152 for French, 950 for Chinese, 884 for German, 432 for Dutch, 368 for Portuguese, 367 for Italian, 285 for Turkish, 266 for Arabic, 116 for Welsh, 53 for Hindi, and 48 for Creole. No filters or language specifications were added. This quick search highlights the tendency to study some languages more than others. For example, in this database, there were only 48 studies about creole languages while there are at least 24 creole languages in the world (Michaelis *et al*., 2013).

This bias in science is not random. It stems from social and geopolitical power dynamics because it is not possible to separate science from the society which produces it.

"[K]nowledge production [...] is a key component of the broader social and political relations in which it occurs" (Boncourt; Ravecca, 2020, p. 95). In parallel, science can be considered a public good (Boulton, 2021) in that the results from scientific studies should be used to the benefit of peoples and should be communicated as such. Consequently, in order for science to reach these objectives, it is necessary that researchers actively pursue projects which contribute to diminishing this bias.

In the case of Linguistics and bilingualism, this phenomenon does not involve only specific varieties of English, Spanish and French. As previously mentioned, the fact that people may use more than one language in their everyday lives is very common. However, these languages are not restricted to high-prestige ones. There are languages in different modalities, such as the many sign languages around the globe, and there are minority languages, which is a broad category. A minority language is used by ethnolinguistic minority groups in the context of a majority language (Montrul, 2015). Minority languages are usually marginalized due to social, political and cultural factors while the majority language is the one used in means of communication and in educational, administrative, and governmental levels. Here one could include indigenous languages, languages from migrant people, creoles and pidgins, and languages at risk of extinction. The first mention of minority languages may be from 1969 (Walsh, 1969).

In 2010, Henrich, Heine, and Norenzayan published a paper promoting a debate on the impact of the reliability of scientific results, specially in the social and psychological sciences, when considering that most of the evidence was based on studies with specific populations. Some of their arguments may be an overextension of study results and were indeed questioned. However, the need to have more diverse samples in social and psychological studies and to test hypotheses with varied populations is real (Leivada *et al*., 2023). There are at least indirect effects from environmental factors, such as socioeconomic status (SES), which has an effect on cognitive development (Engel de Abreu *et al.,* 2015; Noble; McCandliss; Farah, 2007). For example, results from studies with participants from high-SES backgrounds may not generalize to samples from low-SES ones. It may not be wise to simply conclude that results for specific populations will fit perfectly for different ones (Henrich; Heine; Norenzayan, 2010).

In this study, in order to contribute to the body of knowledge about populations and languages of less social prestige, our active choice was to consider a minority language in Brazil. Haitian Creole (HC) is the mother tongue — or one of them — of immigrants from Haiti and from the Dominican Republic. Immigration from Haiti to Brazil started in 2010 because of a catastrophic earthquake. However, since its independence from France in 1804, Haiti has been through political and economic crises involving dictatorships, military occupations, coup d'etats, and natural disasters. Collaboration between Brazil and Haiti has existed since 2004 and has been strengthened in 2010 and 2023. Due to the earthquake of 2010, about 161,000 Haitian people have settled in Brazil. Until now, according to a migration database called SisMigra, 173,493 people entered Brazil legally coming from Haiti and got registered. They have the right to have decent jobs, to access public policies, and to be included in the society and economy.

Other than HC, Brazil is home for many minority languages. The estimate is that there are 250 languages in the country, and they can be categorized as sign languages, indigenous, Creole and African-Brazilian languages, migration languages, and many varieties of Brazilian Portuguese (BP) (IPHAN, 2014). Each of them has a history and demands for protection policies and should be thoroughly described and studied. Here we mention the work done by the Institute for Investigating and Developing Linguistic Politics (IPOL) in mapping and preserving minority languages and offering pedagogical, political, and legal advice. Despite all this linguistic diversity, our interest in HC originated from the similarities between HC and BP. More specifically, the form and their respective meanings of words are similar across this pair of languages due to Latin influences from French. Consequently, HC and BP should have a broad overlap in the form of words in their lexicons. According to Lefebvre (1998), French seems to have supplied most of phonological representations and lexical items in HC while Fongbe and related languages appear to have provided many syntactic and semantic properties found in HC. Thus, because lexical similarities between these languages are apparent, it is innovative to investigate parallel language activation via cognate words here.

Cognate words have similar form and meaning across languages (van Assche *et al*., 2020). Usually there is more overlap in orthographic form than in phonological one, but most of the semantic traits are maintained. For example, *piano* and *piano* are identical cognates in English and BP while *bank* and *banco* are non-identical ones. Words like cognates, homophones, homographs, and homonyms have been used for more than 50 years to study language co-activation. They have been shown to either facilitate or to inhibit language processing of bilingual people whenever they are encountered. This was considered evidence that both languages of the bilingual are activated and influence one another.

Another consequence of the study with cognate words is discussing the architecture of the bilingual lexicon. The most accepted description has been that a bilingual person may have one lexicon for each language, but that they are integrated. It is because of this interconnection of lexicons that cognate effects are seen. Otherwise, these words would be processed no differently than any noncognate ones. Indeed, bilingualism does change native-language reading for example (van Assche *et al.*, 2009). This hypothesis is also extended to the multilingual mind, and many models of bilingual language processing have proposed their versions of this organization.

The Bilingual Interactive Activation Plus (BIA+) model (Dijkstra; van Heuven, 2002) and the DevLexII (Li; Zhao, 2013; Li; Zhao; MacWhinney, 2007) both propose integrated lexicons in different implementations. BIA+ bases the word identification system on bottom-up processes, that is, phoneme and letter identification, which will activate words containing them and inhibit ones which do not. Meanwhile, DevLexII has many self-organizing maps onto which lexical representations are represented according to patterns of activation. These patterns are boosted every time a word is encountered again, and weakened when it is not. BIA+ explains that cognate words receive activation from both languages because of the overlap in form and meaning. DevLexII does not mention cognate words specifically, but one could speculate that the mappings for them might be stronger due to more activation received by the patterns.

In sum, cognate words can demonstrate that the languages of a bilingual person are co-activated. There is evidence from auditory and visual comprehension and from production tasks. For example, Thierry and Wu (2007) showed that when listening to and reading English word pairs, Chinese-English bilinguals were influenced by the repetition of one character in the Chinese counterpart of the word pair. Here the repeated character would only have an impact if Chinese could be accessed via English words, that is, lexicons are probably highly integrated. Dijkstra *et al.* (2010) and Toassi, Mota, and Teixeira (2020) also observed language co-activation in reading tasks, and Li and Gollan (2018a) identified this influence during a naming task.

Recent studies have been focusing on describing more fine-tuned aspects of this language co-activation. For example, Comesaña *et al.* (2015) investigated differences on the cognate effect when these words are identical cognates across languages or when they are non-identical ones. Interest on whether these effects generalize to multilingual people has also been increasing. In this case, Toassi, Mota, and Teixeira (2020) observed that adult multilinguals indeed show cognate facilitation effects. However, there are certainly limits to these influences that may involve proficiency, language use, and environment. For example, while the immigrant bilingual children in Woolpert's (2018) study did not demonstrate cognate effects, the bilingual children from high-SES backgrounds in Arêas da Luz Fontes *et al.*'s (2021) did.

Indeed, language processing is influenced by many factors concerning input characteristics and individual differences. One variable which has been associated with language development and especially reading skills is phonological awareness (PA). Based on the definition by Moojen and Santos (2001), PA is noticing that words are made up of different sounds which may be segmented and/or grouped and is also having the skill of operating with these sounds. Many studies show a positive correlation between measures of PA and reading skills, but some point to causal effects. Paula, Mota, and Keske-Soares (2005) observed that the reading and writing skills of 76,47% of children in the experimental group were improved by an intervention with PA and grapheme-phoneme conversion training in relation to the control group. Santos and Befi-Lopes (2012) saw a moderate and negative correlation between PA and spelling errors. And Rezaei and Mousanezhad Jeddi (2020) noticed a direct and negative effect of PA on reading errors.

These results are especially relevant for developing reading skills. Castles, Rastle, and Nation (2018) summarize the importance of teaching how writing systems represent oral language. Learning the systematic grapheme to phoneme conversion, at least in the early stages of learning to read, is essential for creating stable and consistent lexical representations. The studies carried out by Siegelman *et al*. (2020) illustrate how learning these patterns makes reading more efficient. They observed that children who were more sensitive to the regularity patterns of orthography to phonology had also higher scores in a letter-word identification task in comparison with children who depended more on the imageability of words. When these abilities are precise, it is much easier to read with the goal of acquiring vocabulary and knowledge.

This process seems to happen in both directions. As reading skill improves, vocabulary grows, which contributes to reading skills. Thus, as vocabulary and, consequently, lexical quality — which is information of lexical constituents — are enhanced, the experience a person has with words gets richer and more varied. Extending this idea, one could argue that reading habits are one of the aspects which affect language development. In other words, the more diverse materials and genres a person is exposed to in different language contexts, the more benefited they would be. It seems that the more books a person reads, the vaster is their vocabulary (Pratheeba; Krashen, 2013) and the better is their reading skill (Butler, 2011). In addition, there may be a relationship between reading habits in one language and reading habits in the other one. In this case, frequency of reading in the L1 may lead to similar frequency of reading in the L2 (Camiciottoli, 2001). This influence of reading habits may be stronger for the L2 development of intermediate learners (Artieda, 2007) and seems to be indirect, with L1 reading habits predicting L1 vocabulary size, and the latter predicting L2 vocabulary size (Caylak Toplu; Erten, 2023).

As mentioned above, learning to read depends on decoding and also on other language skills, especially listening comprehension (Gough; Tunmer, 1986; Verhoeven; van Leeuwe, 2012). Similarly, L2 PA basically depends on metalinguistic and language skills (Saiegh-Haddad, 2019). Thus, it is plausible to argue that the development of a second or third language may also rely on opportunities for growing a rich and varied vocabulary. This high-quality input is important for any person learning a language, but is crucial when advanced proficiency levels in this language are necessary for having a job and being a productive part of society. Access to linguistic knowledge and language learning opportunities is essential for people migrating to a country where a different majority language is spoken. In these situations, the goal is not just communicating with other people, but also guaranteeing health, housing, income, and basic civil and human rights. It is relevant to add that socioeconomic status (SES) is associated with cognitive and linguistic development. So, avoiding an impoverished linguistic environment is needed in order to learn the language and to read in that language.

In a majority and minority languages scenario, both of them should be maintained. Verhoeven (2000) commented that a minority L1 may help with the development of the L2. Raudszus, Segers, and Verhoeven (2018) observed that L1 vocabulary had an effect on L2 reading, which suggests that a richer vocabulary in the L1 contributes to learning new words and acquiring an L2. Additionally, a meta-analysis showed that "[s]timulating oral language proficiency in both languages can be a key factor in improving school outcomes of bilingual immigrant background children" (Prevoo *et al.*, 2016, p. 1). With these studies, we understand that both the majority and the minority language should receive attention from researchers, schools, and governments.

Thus, our society is based on activities which require highly-proficient reading and language abilities. This becomes clear when it is revealed that

> low literacy is a major contributor to inequality and increases the likelihood of poor physical and mental health, workplace accidents, misuse of medication, participation in crime, and welfare dependency, all of which also have substantial additional social and economic costs (World Literacy Foundation, 2015) (Castles; Rastle; Nation, 2018, p. 5).

Meanwhile, decoding and reading comprehension are influenced by other language skills in the same language — and this relationship seems to have an impact in a second or third one as well. These effects between skills and languages seem to be indirect, but their repercussions are being studied. So, in an ideal world, the entire population, immigrants included, would have access to rich and varied materials for language learning in order to help developing all languages they have knowledge.

In sum, bilingualism is a common phenomenon in the world and has been investigated in relation to psychological, neurocognitive, social and educational aspects since before the 1970s. This body of research has revealed that bilingualism is complex (Poarch; Bialystok, 2015) and is characterized by diverse contexts and times of acquisition, contexts and frequency of use, and levels of linguistic skill. Thus, it is fruitful to extend studies to varied types of bilingualism and combinations of languages (Henrich; Heine; Norenzayan, 2010). To my knowledge, there are no other psycholinguistic studies about speakers of Haitian Creole and Brazilian Portuguese. This population started migrating to Brazil in 2010 and has needed to learn BP to communicate and to thrive in this society. It is expected that, similarly to bilinguals and multilinguals in other studies (Lemhöfer *et al*., 2008; Marian; Spivey, 2003; Thierry; Wu, 2007), HC-BP bilingual speakers will also display indications that both languages are active during listening and reading in BP. These indications include the cognate facilitation effect, which is the facilitation of the processing of cognate words during comprehension (Bultena; Dijkstra; van Hell, 2014; Dijkstra *et al.*, 2010). For example, the word *piano* will be recognized faster by a BP-English bilingual than a monolingual because it shares orthography, phonology, and meaning across languages.

In addition, word recognition and lexical access are influenced by individual differences, in this case phonological awareness and reading habits. Learning to read and language development as a whole are influenced by metalinguistic knowledge such as phonological awareness (Castles; Rastle; Nation, 2018; Paula; Mota; Keske-Soares, 2005). Vocabulary also benefits from reading habits (Butler, 2011; Pratheeba; Krashen, 2013), and the frequency of reading and materials which are read can influence language development and favor L2 learning (Camiciottoli, 2001). Moreover, encountering words in varied contexts and having a diverse experience with words helps improving lexical knowledge (Taylor; Perfetti, 2016). However, it would not be safe to simply generalize results from one population to another when their realities are dissimilar (Henrich; Heine; Norenzayan, 2010). More specifically, there are no studies investigating bilingual language co-activation involving HC and BP in a scenario of migration. This gap matters because, despite predictions from bilingual language processing models such as the BIA+ (Dijkstra; van Heuven, 2002), Spanish-English migrant children did not show cognate effects (Woolpert, 2018), while BP-English children from a prestigious school did (Arêas da Luz Fontes *et al.,* 2021). With this literature in mind, I identified a gap to be investigated in this dissertation. Thus, the research questions guiding this work were the following:

1. Considering that bilingual and multilingual speakers have their languages co-activated at some level (Lemhöfer *et al*., 2008; Marian; Spivey, 2003; Thierry; Wu, 2007), do HC-speakers show any influences from HC on lexical access during spoken and written comprehension in BP? More specifically, may HC co-activation facilitate BP processing?
2. Do BP phonological awareness and reading habits in BP play a role in comprehension of BP for HC-BP bilinguals?

Research question number 1 originated the general objective of this work, while research question number 2 lead to two specific objectives. The general objective of this dissertation is to investigate the effects of cross-linguistic interaction between HC and BP on lexical access during comprehension of spoken and written BP by native speakers of HC. The specific objectives are two-fold:

1. To investigate whether phonological awareness (PA) in BP influences the processing of BP as an L2 by native speakers of HC and whether it affects language co-activation effects;
2. To investigate whether reading habits (RH) in BP influences the processing of BP as an L2 by native speakers of HC and whether it affects language co-activation effects.

In order to pursue this general objective, I conducted two separate studies: one with children and teenagers — Study 1 —, and the other with adults — Study 2. Study 1 also examined the specific objective number 1, and Study 2 also examined the specific objective number 2. The two studies will be described separately in two different chapters, each consisting of the report on the method, the results and a brief discussion for each study. Therefore, this dissertation is organized as follows: literature review and theoretical background relevant for both studies; report on Study 1; report on Study 2; a general discussion involving results from both studies; an experience report about the data collection process; and final remarks.

2 LITERATURE REVIEW AND THEORETICAL BACKGROUND

2.1 LANGUAGE PROCESSING

2.1.1 Spoken language processing

At least since the end of the nineteenth century (Stäger; Von Jagemann; Joynes; Raddatz, 1887), there has been interest in understanding the processes behind spoken language comprehension and production. This investigation might be as complex as disentangling the many layers of acoustic information in the speech stream in order to perceive linguistically relevant sounds. And studies on speech perception have been carried out at least since 1947, with the book “Visible Speech”, by Potter, Kopp and Green.

Liberman *et al.* (1968) define the sounds of speech as a special code. This concept comes from the fact that slicing the speech stream into independent phonemes is not an executable task. The reason for this is that there is always some overlap between the pieces of this code. The way speech is perceived, a sequence of phonemes is actually restructured into a larger unit, such as a syllable, in order to convey a message. In other words, “[t]he intermixing and overlapping of the acoustic representations of the phonemes to form syllables is the essence of the code” (Liberman *et al*., 1968, p. 23)

Carol Fowler and James Magnuson (2012) emphasize that “[a]coustic speech signals do not consist of sequences of discrete phone-sized segments” (Fowler; Magnuson, 2012, p. 5). They define speech perception as the mapping of acoustic and even visual information onto language forms (Fowler; Magnuson, 2012, p. 3). However, acoustic forms of language are not produced in a clear-cut, completely sequential manner: “the code transmits the phonemes in parallel in the sense that, at every instant, the acoustic signal provides information about more than one phoneme” (Liberman *et al*., 1968, p. 24). When we perceive speech, we are exposed to the speech continuum and its variabilities — which include coarticulation, assimilation and also the talker’s variabilities — and we organize the acoustic information into categories. That is, phonemes are perceived categorically and “continuous variations in the acoustic cue are perceived discontinuously” (Liberman *et al*., 1968, p. 24) although they are not produced with defined boundaries in the speech flow.

An important tool to visualize speech signals and test hypotheses was the Pattern Playback machine, which scans spectrograms and converts them into sounds (Cooper; Liberman; Borst, 1951). It enabled the manipulation of spectrograms for studying the speech stream. This way, researchers could use both auditory (signals) and visual (spectrogram) information to identify acoustic cues and coarticulation contexts (Fowler; Magnuson, 2012). Coarticulation, which is the overlap of articulatory gestures, causes assimilation, when a segment assimilates the value of a neighboring segment's feature. Phenomena such as these lead to variability in speech signals, and the listener has to adjust signals and use cues, such as duration, stress, and context, in order to map them onto existing phonetic categories (Fowler; Magnuson, 2012).

Speakers seem to detect subtle lexical and sublexical acoustic properties of words and parts of words, such as phonotactic constraints, length, metrical and lexical stress, neighborhoods densities, and speaker characteristics (Samuel; Sumner, 2012). Using that information from the speech stream, they adapt or "'tune' their phonemic representations to reflect the incoming speech signal" (Samuel; Sumner, 2012, p. 68). A consequence of this is that multiple representations of lexical candidates can be activated during spoken language processing, and context does not seem to impede this (Samuel; Sumner, 2012, p. 62, 64). This will be important for the discussion on cross-language activation later in Section 2.2.2.

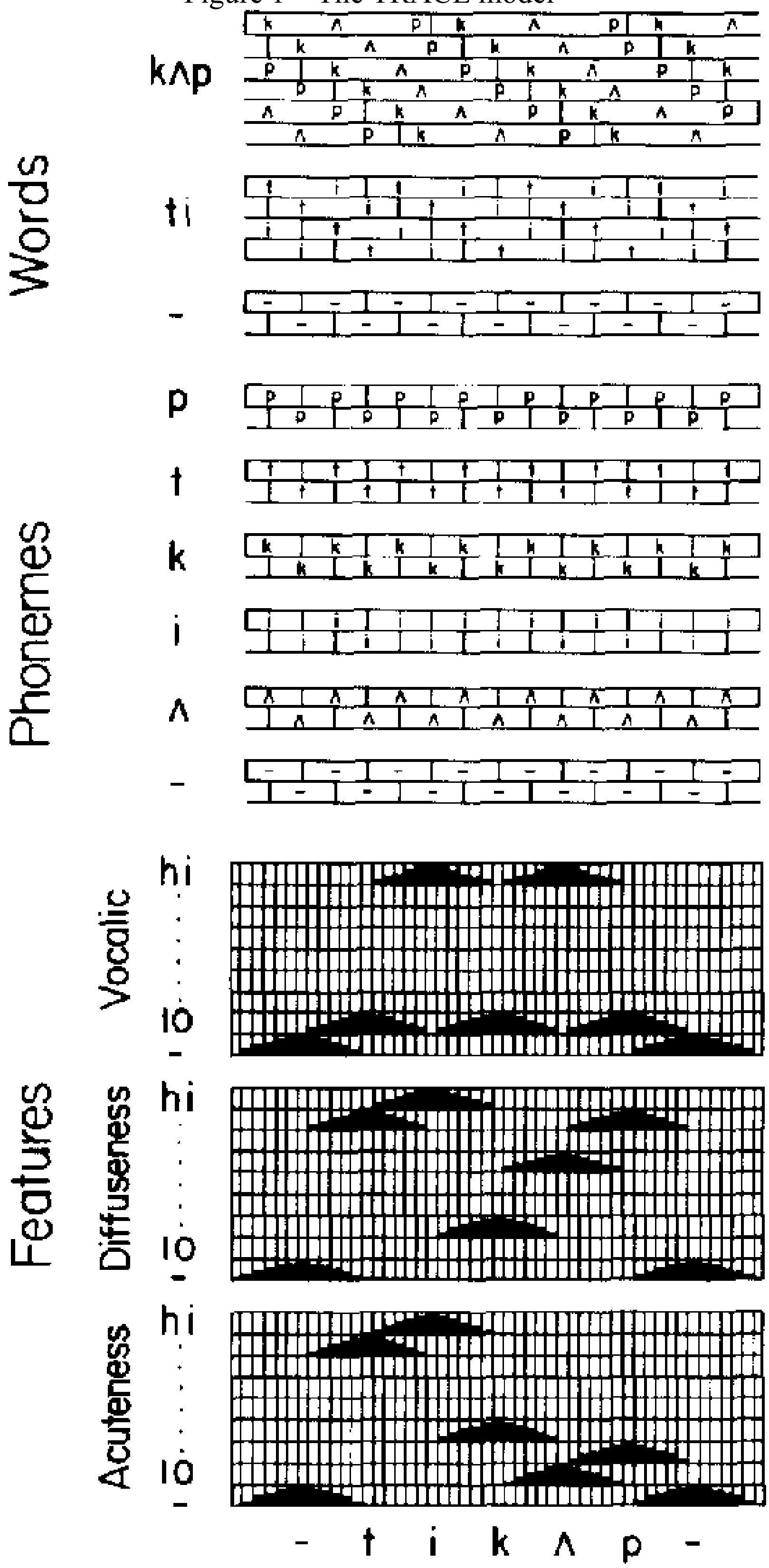
However, speech conveys messages, not just phonemes, and extracting words from the speech stream is virtually a task full of obstacles which listeners and speakers do manage to accomplish. Two characteristics of the speech stream which pose difficulties for recognizing spoken words are the lexical segmentation and the embedding problems (Fowler; Magnuson, 2012, p. 17). The first one comes from the fact that the speech continuum does not have enough cues to clearly define word boundaries, that is, it may be hard to draw the exact moment when a word begins and ends. This leads to the second problem, in that smaller words are embedded into bigger words, such as *cap* in *captain* and *bar* in *barco* (*bar/pub* in *boat*). Since word boundary cues are not clear, the listener may encounter many competing lexical candidates when parsing speech.

A similar situation occurs when perceiving sounds from a non-native language. One of the first studies on non-native speech perception seems to be the one by Horatio Hale, published in 1885, where he addresses the interchangeable use of specific phonemes in certain languages and dialects and "intermediate" phonemes. His remarks make one characteristic of the L2 speech perception very clear: the influence that L1 sound categories play on learning to perceive the new L2 sounds. Bohn (2018, p. 225) explains that this process is guided by "how L2 sounds are mapped unto L1 categories [...] [and by] L2 experience, the age factor and language use patterns". In this case, the age factor is more of an expression of varied experiences during the development of both phonetic systems than of simply age as maturation (Bohn, 2018, p. 214).

The monolingual infant is exposed to the ambient language(s) since before birth and is gathering information on the linguistic sounds to create a "multidimensional perceptual space" which will be tuned and specialized for that language. Thus, when a listener is learning a second language, the sounds of this new language do not fit in the existing perceptual space since it was designed for the native language systems (Bohn, 2018, p. 215). This means that the "attentional preferences" of these listeners should be modified in order that sound contrasts across languages do not become learning difficulties (Bohn, 2018, p. 219).

One of the many models which structured the steps and processes behind speech perception is the TRACE model (McClelland; Elman, 1986), which is a connectionist model which simulates speech perception. It is composed of units which function as detectors of speech characteristics (Figure 1). These units are divided into a level for features, one for phonemes and another for words. The detectors are replicated for many time slices, which is a mechanism able to account for the temporal overlap of gestures (Dahan; Magnuson, 2006, p. 255). These units/detectors interact with each other in excitatory or inhibitory ways, that is, the spread of activation and inhibition is bidirectional, just like in an interactive activation approach (McClelland; Elman, 1986, p. 2). This means that, after a unit receives enough activation, it excites other units which match the input and inhibits units which do not match the input. However, unlike the Interactive Activation model (McClelland; Rumelhart, 1981), which will be presented in Figure 2 in Section 2.1.2, in TRACE units from one level do not inhibit units from other levels (McClelland; Elman, 1986, p. 12). Thus, all active units form the representation candidates — the hypothesis — which match the perceived input.

Figure 1 – The TRACE model



Source: McClelland and Elman (1986, p. 9)

Units in TRACE detect features and phonemes in speech. However, features vary according with the phonemic context they are produced in. TRACE is able to adapt and detect those differences due to the many time slices, which influence one another. Each time slice in every level is actually one processing cycle. This strategy provides an analog to coarticulation in the model (Dahan; Magnuson, 2006). Since previous and subsequent time slices are compared in order to update the activation of units, it may be that incompatible units are activated. When inconsistent units are activated at the same time, they end up inhibiting each other. Thus, TRACE deals with the phoneme and the segmentation problems by creating inhibition within levels when features and phonemes coincide in time (Dahan; Magnuson, 2006).

TRACE I was created for processing real speech, while TRACE II was constructed to accommodate phoneme perception in word context. McClelland and Elman (1986) discuss only TRACE II in more detail in their paper. This version of the model has seven acoustic-phonetic dimensions, which are simplifications of real speech. Each dimension is composed of a value which may range from 1 to 8. The patterns of values across dimensions represent specific phonemes. Speech perception occurs as follows. The input is presented to the initial time slices, when activation values are set to resting state. Activation spreads to units in the initial slice, and values are updated until the cycle is finished. The next cycle begins, and values are updated in relation to the previous time slice. The three levels receive activation simultaneously, and units spread over time due to the many time slices. This enables TRACE to predict temporal dynamics. “TRACE represents time: Words that become activated early in the spoken input have an advantage over the words that become activated later” (Dahan; Magnuson, 2006, p. 257). The same was true for phonemes (McClelland; Elman, 1986, p. 28). TRACE also accounts for lexical effects and their variances in phoneme perception and for phonotactic rule effects (McClelland; Elman, 1986, p. 33).

Thus, phonemes are perceived and categorized as such and reanalyzed in relation to the subsequent ones. During this process, activation spreads to words which are composed of the identified phonemes, and this activation is updated as perception continues. Some of these steps are also proposed to occur in written word processing. However, spoken and written word processing are inherently different from written word processing. Lexical decision tasks performed either with auditory stimuli (spoken words) or visual stimuli (written words) produce variation in response times (RT) which may reflect these processing differences. One of the first studies to test whether modality would influence lexical decision was a semantic priming one by Swinney and colleagues (1979). They invited 24 participants — there was no mention of languages spoken — to complete a primed lexical decision task using auditory stimuli as primes and visual stimuli as targets. Mean RT for unprimed words was 886 ms and for primed words was 803 ms; mean RT for unprimed pseudowords was 1053 ms and for primed pseudowords was 1087 ms (Swinney *et al*., 1979, p. 160).

However, Swinney *et al.* (1979) did not actually compare performance in the lexical decision task across modalities. With this specific goal in mind, Tucker *et al*. (2019) conducted a megastudy contrasting data on visual and auditory lexical decisions. They comment that “auditory experiments are labor-intensive” (Tucker *et al*., 2019, p. 1188) and that there are few studies setting lexical decision results across modalities side by side probably due to this arduousness. For this megastudy, 231 adult monolingual native English speakers completed an auditory and a visual lexical decision task. Some of the analyses Tucker *et al*. (2019) report show that participants responded faster to visual than auditory stimuli. From the information available in the paper, mean RT for spoken words was 940 ms and mean RT for written words was 759 ms (Tucker *et al.*, 2019, p. 1200).

Similarly, Zunini *et al*. (2020) compared results from auditory and visual lexical decision tasks; however, they included an audiovisual version of the task and registered behavioral and neurophysiological data. The 22 adult native Spanish speakers — there was no mention of whether participants were monolingual or bilingual — who participated also responded faster to visual than auditory stimuli. At this moment, I will be commenting only on the behavioral data. From Figure 1 in Zunini *et al.* (2020), in the visual task, mean RT for words was about 820 ms and for pseudowords was about 980 ms, while in the auditory task, mean RT for words was about 1200 ms and for pseudowords was about 1300 ms (Zunini *et al*., 2020, p. 5).

In the next section, I will present a brief account of the factors which influence written language processing.

2.1.2 Written language processing

One of the first studies on visual word processing comes from ophthalmology. It was published in 1938, and only the first page was available online. All I could gather about this work is that Thomas Eames tested 40 children with reading difficulties and 50 children without any difficulties and observed that the first group recognized words more slowly than the second one. The latencies are indicating individual differences across groups, and since then response times have been an important measure for reading. In 1957, Adis-Castro and Postman were interested in testing two methods for studying visual word recognition. Independently of method, they observed the now classic effects for frequency and for length: faster responses for higher-frequency words in comparison with lower-frequency ones and slower responses for longer words in comparison with shorter ones. These studies illustrate a few characteristics of the reading process which have been observed since 1938.

These examples show the effects of input characteristics on reading. The speed and the precision with which the reader recognizes a word are affected by lexical and sublexical properties. Examples of lexical properties are word frequency, length, orthographic and phonological similarity. Examples of sublexical ones are regularity, consistency, orthographic and phonological neighborhood densities. I will present a brief overview of their effects.

Word frequency is probably the most mentioned feature to control for in studies on word reading. It has also a robust effect in written word recognition. It impacts both speed and accuracy of responses: words which occur more frequently are responded to faster and more accurately than words which occur less frequently (Yap; Balota, 2015). Since it is so prevalent, this feature was added to word recognition models and is part of the predictions of these models. For example, as I will explain in more detail later, the Interactive Activation model (McClelland; Rumelhart, 1981) implements the word frequency effect as various resting-activation levels. In other words, the amount of baseline activation which a word has, that is, when no stimuli from the input is being received is higher when the word is highly frequent. Thus, this word will require less information from the input in order to be activated in comparison with low-frequency words. A different perspective on frequency is familiarity, which varies according the person's linguistic experience. Lexical familiarity may be considered an individual differences variable (Lewellen *et al.*, 1993) and is related to the encounters a person has had with a word. Familiarity is correlated with frequency, but a familiar word is not necessarily a frequent one (Tanaka-Ishii; Terada, 2011).

Word length is related to the number of letters a word has. It is also important in written word recognition, but with the number of phonemes. In a simple description, words with more letters will take longer to be recognized than words with less letters. Length effects are better explained by models which rely on serial processing, such as the Dual-Route Cascaded model (DRC) (Coltheart *et al*., 2001), instead of by models based on parallel processing, such as the Triangle model (Harm; Seidenberg, 2004; Seidenberg; McClelland, 1989), and are larger for nonwords than real words (Yap; Balota, 2015).

Yap and Balota (2015, p. 34) define regularity as "the most statistically reliable spelling-to-sound correspondence rules in the language" and consistency as "the extent to which a word is pronounced like similarly spelled words". The effect generally seen here is that irregular and inconsistent words are responded to slower than regular and consistent ones. For example, in English, *hint*, *mint*, *tint* and *lint* follow a regular pronunciation of *i* in the *-int* context. However, *pint* is irregular because here the *i* is pronounced differently. Also, pint is inconsistent in relation to the previous similarly spelled words. Knowing these patterns and their exceptions is important when learning to read in an alphabetic system.

Other sublexical properties are orthographic and phonological neighborhood densities. That is, a word may have neighbors via orthography and/or via phonology. A word neighbor can be a word with similar form in that one letter or phoneme was transposed, replaced, removed or added (Perea, 2015, p. 79). There can also be syllabic neighbors. Due to the visual and auditory modalities, orthographic and phonological neighbors cause different effects and they also interact with frequency. For example, words with a large phonological neighborhood size were faster to be visually recognized than words with small phonological neighborhoods (Yates; Locker Jr; Simpson, 2004). The presence of a higher-frequency neighbor in the orthographic neighborhood hinders access to a word (Grainger *et al*., 1989). Visual recognition of low-frequency words is facilitated when they have a denser orthographic neighborhood (Andrews, 1989, 1992). In addition, different task demands produce diverse effects. Large neighborhood sizes create facilitation effects in lexical decision tasks while producing inhibition effects in word identification tasks.

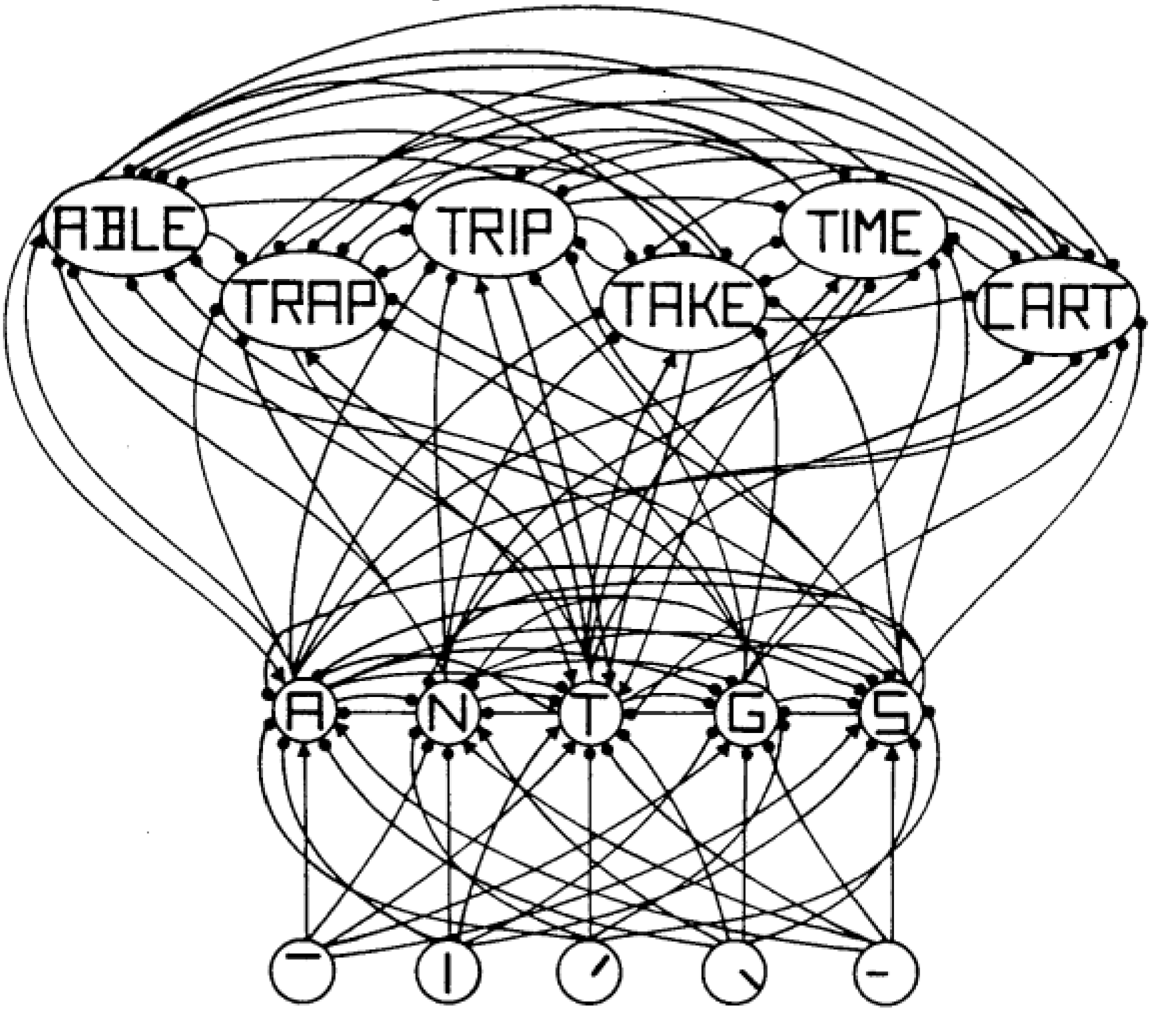
A key lexical property for this study is form similarity: orthographic and phonological similarity. It may be considered a priming effect instead of a lexical one. Phonologically similar words might facilitate (Lukatela; Turvey, 2000) or hinder (Lemhöfer; Huestegge; Mulder, 2018) word recognition, and comparable phenomena happens for orthographically similar words: facilitation (De Groot; Nas, 1991, within and between languages) or inhibition (Brenders; van Hell; Dijkstra, 2011). These effects can even be observed across modalities as mentioned above (Swinney *et al*., 1979) and between languages, which will be the focus of Section 2.2.2.

As commented above, all these properties were added to models of word recognition in order to propose how this process works and what systems are involved. Models also offer predictions on the phenomena and may simulate, in the case of computational ones, the steps towards word reading.

The Interactive Activation model (IA) (McClelland; Rumelhart, 1981) is a connectionist model of print perception. It was based on features from the pandemonium (Selfridge, 1959) and the logogen (Morton, 1969) models. The pandemonium model inspired the specific detectors for letter features, which had a neurobiological basis from studies showing that there are specific neurons which identify very basic lines (Cortese; Balota, 2012). The logogen model inspired the detectors in the IA model. Logogens would detect sensory information of a word and accumulate it until reaching a threshold, when the word would become available for processing (Braze; Gong, 2018). The amount of activation to engage a word would vary according to word frequency: higher-frequency words would need less activation because they have higher resting-level activation in comparison with lower-frequency words (Cortese; Balota, 2012).

As may be seen in Figure 2, the IA model is composed of 3 levels of detectors: one for features of letters, such as vertical, horizontal and diagonal lines, one for letters, and one for words. Words are represented in the model by nodes. The detectors of features identify them in the input and send activation to letters which match with the detected features and inhibit the letters which do not match with the input. Letters receive this excitation and send it to words which have those letters while inhibiting words which do not have them. The activation spread through these links gets accumulated over time and flows in both directions between levels. That is, features are activated and send activation to letters; letters are activated and send activation back to features and forward to words. Thus, there are excitatory and inhibitory links among nodes within a level and between nodes across adjacent levels. Also, the IA model is able to account for the word superiority effect, which is the fact that letters are more easily recognized when embedded in words than in nonwords.

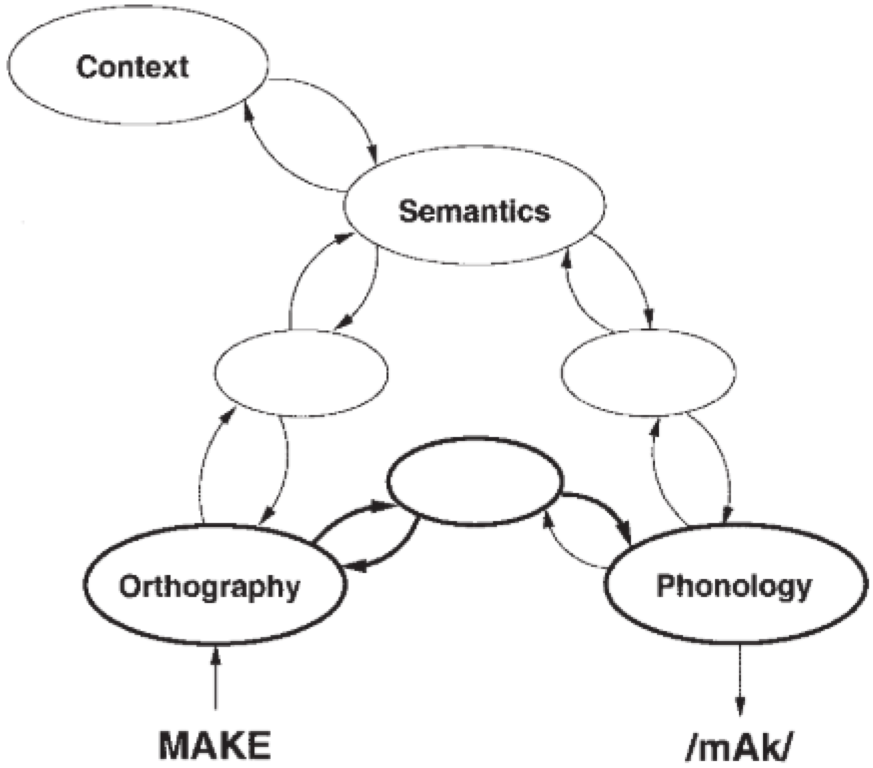
Figure 2 – The IA model



Source: McClelland and Rumelhart (1981, p. 380)

Another relevant model about the processes involved in reading is the Triangle model (Harm; Seidenberg, 2004; Seidenberg; McClelland, 1989), which is a connectionist model for describing word reading from print to meaning. The model assumes that a word is an "intersection of phonological, semantic and orthographic information" (Braze; Gong, 2018, p. 279), which is in accordance with one of Charles Perfetti's (2007) proposals in the Lexical Quality Hypothesis. A key feature of the model, which is represented by its triangular shape (Figure 3), is the division of labor between the direct pathway from orthography to meaning and the mediated pathway from orthography, to phonology and then to meaning. This means that both pathways are at cooperative work at the same time and that meanings are activated because of this cooperation.

Figure 3 – The Triangle Model



Source: Harm and Seidenberg (2004, p. 663)

The model has non-symbolic lexical representations which are composed of sets of distributed units of the three codes: orthography, phonology and semantics (Braze; Gong, 2018; Harm; Seidenberg, 2004). Thus, in this model, the meaning is not "accessed" as in a localist approach, it is computed from the activation patterns of sets of distributed units. Also, the model is equipped with a backpropagation algorithm which allows it to learn in a supervised manner. This feature is used to pretrain the model on the connections from phonology to semantics, emulating a child learning their spoken native language. Then, in order to learn to read, the model uses statistical learning mechanisms. Whichever pathway achieves the goal faster or more efficiently is the go-to pathway in that case. The direct pathway processes words with more consistent pronunciations more easily while the mediated pathway processes less consistent words more adequately; in the case of homophones and homographs, both pathways contribute (Harm; Seidenberg, 2004, p. 672).

In sum, the Triangle model offers relevant mechanisms, such as the division of labor, for describing the processes behind reading words in search of their meaning. In the next sections I will present a summary of the literature about learning to read, the stages of learning to read and constructs associated with it, for example phonological awareness.

2.1.3 Learning to read

Learning to read is dependent on the type of writing system the language uses. Writing systems can be logographies, syllabaries or alphabets (Kessler; Treiman, 2015). English, Brazilian Portuguese and Haitian Creole use alphabetic writing systems to represent spoken language. Alphabets use letters to indicate sounds; more specifically, graphemes represent phonemes, that is, "letters code phonological information" from the spoken language (Nation, 2012, p. 206). Learning to read is a hard task because it is not an innate ability, such as spoken or signed language, and it requires understanding rules that the code uses to embody linguistic information (Kessler; Treiman, 2015, p. 23-24). Thus, it is necessary to learn the patterns of grapheme-to-phoneme conversion in order to decode words and read. However, "visual word recognition alone is not sufficient to guarantee successful reading comprehension" (Nation, 2012), which does not mean that the code should not be explicitly taught.

The development of reading as a skill depends on many factors, including letter knowledge, grapheme-phoneme conversion rules and metalinguistic knowledge. One of these is phonological awareness (PA), which can be defined as "the capacity to represent phonological properties consciously" (Morais *et al.*, 1998, p. 66). Some researchers extend that definition and state that it "refers to an awareness of and access to the sound structure of one's oral language" (Wagner; Piasta; Torgesen, 2006, p. 1114). In this study, I will use the definition by Moojen and Santos (2001, p. 751, my translation) that PA is

> the awareness that words are composed of different sounds or group of sounds and that they may be segmented into smaller units. Phonological awareness involves not only the ability of thinking (noticing and comparing) but also the ability of operating with these syllables or phonemes (counting, segmenting, binding, adding, removing, replacing, transposing).

Nation (2012) explains that PA and letter knowledge have an impact in predicting reading skills (Byrne; Fielding-Barnsley, 1989; Muter *et al*., 2004; Hatcher; Hulme; Ellis, 1994). More specifically, letter knowledge influences later PA, which in turn directly impacts word-level reading (Wagner; Piasta; Torgesen, 2006, p. 1117). In fact, there is evidence from a 5-year longitudinal study pointing to a causal role of PA in learning to read (Wagner *et al*., 1997).

Moreover, the influence of word reading skill and PA seems to be mutual (Wagner; Piasta; Torgesen, 2006, p. 1122). This relationship is regulated by many aspects, but it is not clear where it originates from. One of these aspects is phonological training, which seems to produce better results to reading skills than phonemic awareness training by itself, especially with at-risk students (Torgesen *et al*., 1999, p. 1119). A second one is the fact that vocabulary and reading comprehension mutually influence each other. A third explanation for the relationship between word reading and phonological awareness would be that conceptual knowledge or general verbal ability is mediating it. Finally, it could be that metalinguistic awareness is related to vocabulary and reading comprehension. Independently of the source of the relationship between word reading and phonological awareness, they do impact reading development.

Learning to read requires knowledge, skills, instruction, and exposure to language. Ehri (2015) presents strategies used by children when learning to read words in a phonographic system and the phases they go through when developing their reading abilities. The first strategy is decoding, which involves grapheme-phoneme conversion. Decoding requires having the knowledge about the associations between letters and sounds of a language and the skill to identify these relations and maintain them in mind in order to access the word and its meaning. Decoding is especially important for testing possible pronunciations of known words whose written form is unknown and for dealing with the spelling variability of words. After learning to recognize graphemes, it is time to recognize larger units such as syllables, rimes, and morphemes, which makes decoding faster since these patterns become more regular than individual graphemes.

The second way of reading words is through analogy. Analogizing or associating known patterns and units with unknown strings of letters is different from decoding them, but it is an effective strategy too (Ehri, 2015). However, being able to decode seems to be important to use analogies more efficiently or with benefits. The third strategy is prediction, which depends on the reader's prior knowledge and the available context information. It is possible to decode the first graphemes of a word and then try and predict it. Although prediction seems to compensate for some lack of decoding ability, the latter is shown to improve the former. Finally, the last way of reading words is relying on sight words. Once the reader has enough practice in reading words, they may recover them from memory by sight. It is a faster and automatic process since words are considered whole units instead of sets of graphemes which need to be decoded. Ehri (2015) argues that this is the most efficient reading strategy. However, sight words must be learned, and grapheme-phoneme connections are key. Learning these links is important for binding spellings, pronunciations, and meanings.

Moreover, Ehri (2015) presents a theory of phases of development of word reading. The first one is the prealphabetic phase, in which children are not aware of the grapheme phoneme connections. They tend to use random letters, or invented ones, to represent words and apply contextual cues to recognize them. However, they are still not aware of letter sounds or names. The second phase is the partial alphabetic one, when there is some knowledge of the links between graphemes and phonemes, but it is still incomplete. At this stage, children are not able to decode or to memorize words yet. They use cues from context and partial information from letters to guess words. The third phase is the full alphabetic one, in which many connections between graphemes and phonemes are recognized and sight words are being formed. At this stage, children are more efficient and accurate in creating the links between sounds and letters and consequently storing sight words. The last phase is the consolidated alphabetic one, when children are able to decode multisyllabic words using memorized letter patterns, such as morphemes, and sight words. Ehri (2015) also mentions the importance of experience with words in sentential and textual context for vocabulary growth and fluency. It is key to gathering syntactic and semantic knowledge and to understanding function words and context-dependent words.

Ehri (2015) presents one perspective on the stages of learning to read. A distinct one is the one by Frith (1985). Her model was based on the cognitive developmental theory by Marsh and colleagues (1977, 1980, 1981, 1983). Their theory proposes four strategies for learning new words. The first is rote learning, which involves guessing the word from contextual cues. The second is related to using letters as cues for guessing words. The third is sequential decoding, that is, grapheme-phoneme conversion. And the fourth is hierarchical decoding, which takes into consideration the context of every phoneme and uses analogy to learn new words. Inspired by these four strategies, Frith (1985) suggested a new model of the stages in reading development.

Frith (1985) introduced three strategies or phases. First, the logographic phase replaces the rote learning strategy mentioned above. It is the visual and symbolic phase. Children will use their memory skills and will focus their attention on familiar words. The strategy at this stage is to concentrate on striking visual graphic features of words and contextual or pragmatic cues. Children may construct a sight vocabulary of considerable size. Second, the alphabetic phase is equivalent to the strategy of using letters to guess words in Marsh and colleagues theory. This is the phonological phase, when it is recommended to add phonics instruction. It may be that children are mixing sight words and cannot tell them apart. The strategy here is to focus on individual phonemes and graphemes in order to (correctly or not) pronounce new words. This is the stage when phoneme awareness appears. Third, the orthographic phase encompasses the last two strategies in the theory by Marsh and colleagues. This is the abstract stage, when visual and phonological cues are not vital anymore. Children are able to read words without converting every grapheme into phonemes. Now, they consider orthographic units instead of individual phonemes. These bigger units may be morphemes or even syllables. After this stage, written language may be considered a system independent of spoken language. Ehri's (2015) and Frith's (1985) models suggest two different points of view on the steps of how to learn to read.

The debate is not concluded since there is evidence that these stages may be more blurred than first thought. For example, analogies are not used only at later stages of learning to read. Actually, children make analogies since the beginning of reading development (Nation, 2012, p. 208). This suggests that this process is not defined by clear-cut stages and that children develop their sensitivity to the mapping of phonemes to graphemes gradually according to their knowledge of this code.

In alphabetic orthographies, graphemes may represent phonemes with diverse patterns. The consistency of these patterns varies among different orthographies. For example, the English orthography is less consistent than the Italian one. Gradually noticing these patterns is a part of a statistical learning process (Nation, 2012, p. 210) and can help reading development. The studies by Siegelman *et al*. (2020) illustrate the positive impact of noticing regularities for reading. In study 1, 123 third- and fourth-graders completed a word naming task in which word frequency and imageability were manipulated. Also, the probability of vowel pronunciation was calculated as the surprisal value and manipulated in the task. For example, the pronunciation of i in pint has a higher surprisal value than in mint. Participants also completed 4 reading skills tests. Results showed that children who were more sensitive to the regularity patterns of orthography to phonology had higher scores in the letter-word identification task than children who relied more on word imageability. In study 2, 282 children went through the same word naming task but only 3 reading skill tests, achieving similar results. "In both studies greater sensitivity to O-P regularity and lower sensitivity to imageability were associated with higher reading skills" (Siegelman *et al*., 2020, p. 8). It seems that strengthening associations between lexical representations may improve the effectiveness of the statistical learning mechanism (Siegelman *et al*., 2020, p. 11).

Discussion on the process of learning and how to teach how to read is extensive. Castles, Rastle, and Nation (2018) summarize how children learn to read and become skilled readers and explain the importance of phonics instruction and what else should be taught and how to do so. First, the authors emphasize that reading is complex and differs across writing systems. Alphabetic systems represent languages through graphemes, which are letters representing phonemes. This knowledge is called phonemic awareness. Children go through stages of acquiring decoding skills, developing spelling skills and being able to use them consistently. These characteristics of writing systems and the process of learning to read imply that teaching grapheme-phoneme conversion systematically is important for children to develop reading skills, at least in the early stages. The authors mention that teaching sight words along with phonics instruction has a positive impact for learners of languages with deep orthographies. Second, evidence showing that even adult skilled readers use their decoding skills routinely is presented (Rayner *et al.*, 2016). Skilled readers seem to use both the phonological route, for regular and new words, and the lexical route, for irregular and familiar words, during reading. Thus, decoding skills help expand the vocabulary together with exposure to reading materials. Having knowledge about words and having experience with reading them makes their recognition and comprehension more efficient. Morphological knowledge also impacts reading. Then, children should be taught sight words, should receive morphology and phonics instruction, and should be motivated to read. Finally, reading comprehension is influenced by many factors, from decoding skills and vocabulary and background knowledge to non-linguistic cognitive resources and processes, such as attention, inferring, comprehension monitoring, and (working) memory. Thus, comprehension strategies should be taught alongside specific knowledge.

Now we return to the topic of skills which influence reading, Toffoli and Lamprecht (2008) investigated an audio-verbal skills stimulation program and its effect on phonological awareness. Brazilian Portuguese-speaking first-grade children, potentially at a presyllabic stage, participated in the study and completed tasks from CONFIAS3. Despite the effect of the program across experimental and control groups not being significant, the authors explain that there was a significant, strong and positive correlation between participating in the program and performance in tasks tapping the phoneme level. They suggest that there may be an association between phonemic awareness and speech perception. The authors use Moojen and Santos' (2001) definition of phonological awareness, which states that it is the metalinguistic knowledge that words are composed of sounds and may be divided into smaller units and the set of metalinguistic skills that allows one to manipulate those sounds and syllables.

Most of the evidence of the relationship between PA and reading comes from correlational studies. For example, Santos and Befi-Lopes (2012) present a correlational study. They tested whether vocabulary, PA and rapid naming would predict orthographic accuracy in word writing and quality of narrative writing. Participants were 82 4th graders from public and private schools in Brazil. Results for PA showed that PA measures were negatively correlated with error rates in the word writing task and positively correlated with the quality of narrative writing.

Similarly, Rezaei and Mousanezhad Jeddi (2020) were interested in the relationship among attentional components, working memory, PA and reading. They invited 259 children in elementary school, who completed one test for each of the constructs of interest. It was not clear whether the reading test asked participants to read words and sentences out loud, but since error rates were computed, I assumed it was a reading out loud test. Results showed a negative correlation between error rates and PA and a positive correlation between PA and WM. Subscales of attention also presented positive correlations with PA. This relationship is also observed in neurophysiological studies. Frost *et al*. (2009) present behavioral and neurocognitive results pointing in this direction. More specifically, PA measures were positively associated with the activation of brain areas involved in speech when participants were exposed to print.

Longitudinal studies may indicate whether a correlation persists over the time and/or predicts other relationships or measures. Muter *et al*. (1998) investigated how rhyming, segmentation and letter name knowledge influence early reading skills. They conducted a longitudinal study in which at Time 1 38 children were preschoolers, at Time 2 and 3 they were already in school. PA was operationalized as rhyming and segmentation skills. Results showed that PA improved as children advanced in school and that segmentation skill may be more associated with phonemic awareness while rhyming skill may be tapping more general PA. In the first year of school, only segmentation and letter name knowledge predicted early reading and spelling skills. In the second year, only vocabulary predicted reading skill, but rhyming did predict reading and spelling at Time 3. Muter *et al*. (1998) highlight that PA predicted spelling ability throughout both years. This was not an intervention study, but it revealed that PA measures are able to predict skills associated with reading in the first years of school.

Now the study by Paula, Mota, and Keske-Soares (2005) was an intervention one. The experimental group (N = 17) received training on PA and on grapheme-phoneme conversion; the control group (N = 12) did not receive any training; and the alphabetic group (N = 16) was composed of children who could read already. Pretests and post-tests included receptive and expressive language, word and pseudoword reading and writing, and a battery of PA tests. All groups completed the pretest, and only the experimental group and the control group complete the post-tests. Results showed that the control group showed no significant difference in performance in pre- and post-tests while the experimental group did present positive significant differences. In addition, the experimental and the alphabetic group had indicated significant differences during pretests, which became non-significant when comparing the experimental group's post-test with the alphabetic group's pretest. Paula, Mota, and Keske-Soares (2005) conclude that the intervention with PA and grapheme-phoneme conversion training has helped the experimental group improve their decoding skills.

Amorim *et al.* (2020) also conducted an intervention study. They investigated the impact of an educational game called EscriboPlay on word reading and writing of 749 children from 17 private schools. All children completed pretests, which evaluated PA and word and pseudoword reading and writing. Then, they were randomly assigned to the experimental group or the control group. For 10 weeks, the experimental group played with the games in EscriboPlay at school with guidance from teachers. Children in the control group followed regular schooling classes. All children completed post-tests as well, the same ones used for pretests. Results showed significant differences for the experimental group in reading and writing measures. The experimental group had 68% more accuracy in the reading test and 48% more in the writing test in comparison with the control group. Amorim *et al.* (2020) conclude that the PA games together with teacher instruction improved reading and writing skills for the experimental group.

In sum, reading development depends on knowledge of the language, instruction, and exposure to language so that skills may be improved. In the first stages of learning to read, it is essential to practice word decoding because, in an alphabetic writing system, graphemes represent phonemes. Paired with that, reading development also requires exposure to texts and instruction on morphology and sight words. Word recognition and reading are complex processes which become more intricate when considering bilingual readers.

Similar to monolingual reading development, bilingual reading is also affected by knowledge, skills, instruction, and exposure to language. In this sense, Saiegh-Haddad (2019) states that phonological awareness is not independent of language and linguistic experiences, but that it is in fact influenced by two linguistic factors: L2 oral proficiency and distance between L1 and L2. She explains that phonological awareness is both the knowledge that words may be segmented into sounds and the skill to use them. However, she proposes that L2 phonological awareness be composed of a metalinguistic component and a linguistic one. The first one is L2 phonological awareness, and the second one is accurate L2 lexical representations. She explains that, since L1 phonological awareness is separate from L2 phonological awareness, the latter should be taught to children learning to read in a L2. Also, the way phonological information, from all linguistic levels, is represented in the mind can impact phonological awareness. Then, the L2 phonological representations should be precise and detailed. Therefore, the author posits that the improvement of L2 phonological awareness may be affected by L2 linguistic knowledge acquired orally and by structural phonological differences across languages, from phonemes to syllables.

On the topic of learning to read in a second language, Verhoeven (2000) highlights the components of L2 reading and spelling in children learning to read. He explains that the influences of L1 phonological awareness, word decoding skill and vocabulary knowledge are investigated both in L1 and in L2 development and proposes that the strategies used in L1 and L2 reading should be studied. In this longitudinal study, 1,812 children native speakers of Dutch and 331 minority children acquiring Dutch as a L2 completed 8 types of tests distributed throughout grades 1 and 2. These tests tapped word decoding, phoneme and grapheme knowledge, vocabulary, and reading comprehension. The results showed that minority children kept up with their native counterparts on word decoding, but lagged behind on word spelling, phonemic segmentation, L2 orthographic knowledge, vocabulary size, and reading comprehension. Verhoeven (2000, p. 326) concludes that metalinguistic knowledge and oral language experience are necessary to improve word reading, since the decoding abilities in the L2 of the minority group was similar to the native group after 2 years at school, and reading comprehension and that a minority L1 may help with the development of a L2.

Piper, Bulat, and Johnston (2015) present results from two 2-year long randomized intervention programs for improving literacy acquisition in Kenya. The programs were conducted in 847 public and private schools of the country. More than 4,000 children from 1st and 2nd grades participated and were tested 4 times over 2 years. Teachers assigned randomly to experimental groups received training. Students randomly assigned to experimental groups went through the intervention and the testing. The randomization was followed by delayed start of the intervention for control groups. In one of the programs, children were tested in English and Kiswahili; in the other they were tested in Haitian Creole and French. Results showed that both programs had positive effects in reading skills and reading foundation skills. In one of the programs, children showed signs of transference of literacy skills from English to Kiswashili; in the other, "instruction in the mother tongue of Haitian Creole led to gains in French on skills that were not explicitly taught" (Piper; Bulat; Johnston, 2015). These results are especially important for populations in social contexts which lack access to resources and which go through economic struggles. Piper, Bulat and Johnston (2015) mention that most of the teachers in these programs had no previous teacher training and that 40% of all of them had not even completed secondary school. The fact that teacher education was deficient is very concerning for the students' literacy and academic achievements. According to a meta-analysis (Didion;

Toste; Filderman, 2020), on average, professional development of teachers has a significant, moderate, and positive impact on students' reading achievement. In the case of the study by Piper, Bulat, and Johnston (2015), it might be that students are trying to use all relevant knowledge from one language to help with the other even if not consciously.

Considering the role of metalinguistic skills, Verhoeven (2007) investigated the development of L1 and L2 skills of a group of minority children and its relationship with phonological awareness. He explains that bilingual experience should promote an increase in metalinguistic awareness. Also, instruction in one language may be transferred to the other if there is enough exposure to both languages as there are language non-specific skills. Considering that the bilingual lexicon is integrated and that languages are simultaneously activated, the bilingual has to control for interlingual and intralingual processes. Thus, bilingual children should present higher levels of phonological awareness due to their language experience. Seventy-five 5-year-old Turkish speaking children learning Dutch as a L2 participated in the 1-year-long study and completed 7 types of tests which tapped language proficiency, auditory perception, vocabulary, comprehension, imitation, and phonological awareness. Results showed that the children relied more on their L1 during that year; however, L2 proficiency did show improvement. The children's performance in the L2 indicated that they transferred skills from their L1. Moreover, there was an increase in the phonological awareness levels of the minority children group at the end of the year, which suggests that it develops only during kindergarten. Although receptive skills improved, production skills lagged behind; increasing student participation in class may solve this, especially in bilingual education and emergent bilingualism.

Reading in the first and second languages must be based on similar mechanisms. Thus, Verhoeven and van Leeuwe (2012) investigated the possibility that the Simple View of Reading (SVR) (Hoover; Gough, 1990) may be applied to second language learning. The SVR is based on word decoding and oral language skill. Proficient word decoding skill is necessary for efficient word reading and comprehension. Similarly, high listening comprehension ability is essential for a rich vocabulary and reading comprehension skill.

Considering that L1 and L2 abilities develop in parallel and that they may be transferred across languages, the authors tested SVR predictions with a group of minority children. Participants were children who were native speakers of Dutch (N = 1,293) and children who learned Dutch as a L2 (N = 394). The study followed participants from grade 1 through grade 6 and included tests on word decoding speed, listening comprehension, and reading comprehension. Results indicated that performance in all tasks improved across grade for all children. Also, across years, children achieved comparable levels of word decoding skill. However, children learning Dutch as a L2 lagged behind on listening and reading comprehension. Reading comprehension could be strongly predicted both by word decoding and listening comprehension in the first year; in the last year of the study, the predictive power decreased especially for word decoding. Thus, listening comprehension influenced reading comprehension more intensely in grades 5 and 6 probably because word decoding was efficient.

Following Verhoeven and van Leeuwe (2012), Verhoeven, Voeten, and Vermeer (2019) examined how lexical quality influenced the SVR. The Lexical Quality Hypothesis (LQH) (Perfetti, 2007) states that detailed information of lexical constituents, such as orthography, phonology, semantics, and morpho-syntax, contribute to efficient word decoding and, consequently, freeing higher cognitive resources to reading comprehension. A total of 1029 children participated in the longitudinal study, which investigated kindergarten and grades 1 and 2; there were 566 native speakers of Dutch and 463 learners of Dutch as a L2. The children completed tests tapping speech decoding, morphological knowledge, receptive vocabulary, and listening and reading comprehension. Results indicated that all children performed equally in the word decoding tasks. There were moderate differences in speech decoding, and listening and reading comprehension, and large differences in morphological knowledge and receptive vocabulary. The latter variables, measured two years before, predicted reading comprehension directly, confirming the importance of lexical quality. L2 learners still lagged behind L1 learners in all tasks except word decoding, but the predictors were relevant for both groups.

In sum, bilingual reading development involves metalinguistic and linguistic knowledge, and L2 oral language exposure. More specifically, L2 phonological awareness is dependent on accurate L2 lexical representations. In addition, L1 and L2 word decoding skills are the ones both monolingual and bilingual children master first. In the case of bilingual children learning to read in the L2, some of their skills from the L1 may be transferred and may help developing L2 abilities. In the next section, I will present relevant aspects of the literature about bilingualism.

2.2 BILINGUALISM

The first mention of bilingualism I could find was in a paper entitled “The Bilingual-Biracial Problem of Our Border States”, by John D. Fitz-Gerald from 1921. His discussion is much more inclined towards politics and maybe education than linguistics. He concerned himself with the fact that New Mexico, in the young United States of America, had people speaking mainly Spanish who were eager to learn English. He was in favor of the inclusion of English, as well as Spanish, in schools of the region.

The first mention of bilingualism in a more linguistic perspective I could find is in the work of Frank Smith, published in January 1923. It is a paper entitled “Bilingualism and Mental Development” in which he reports two studies he carried out in Wales. The first one, briefly explained, was a cross-sectional study in which monolingual and bilingual students completed 5 reading and writing tests. At the time, they were called monoglots and bilinguists. Smith observed that, in general, monolinguals had better performance when compared with bilinguals. The second study he reports was a longitudinal one, lasting 3 years, with 4 reading and writing tests. In the free composition, mutilated passage and analogies tests, monolinguals performed better than bilinguals; only in the word building test monolinguals and bilinguals had similar scores. Smith concludes that bilingualism may be the reserve of an advantage since he observed monolingual children performing consistently better in those tests.

A few months later, in July 1923, D. J. Saer published the paper “The Effect of Bilingualism on Intelligence”. Saer reports 4-year long longitudinal studies about the effects of bilingualism in children and adults. First, he tested a total of 1400 monolingual and bilingual children from urban and rural areas of Wales for mental age, dextrality, vocabulary, rhythm, and writing skills. It seems every task was carried out both in English and in Welsh for bilinguals. Then, he describes two studies which took place in 1921 and 1922 with university students. In 1921, he devised an intelligence test and asked 278 adults from varied backgrounds to answer it. For the next year, he modified the test and asked 333 adults to answer it. Saer concludes that monolinguals have an advantage over bilinguals even in the academic setting and that bilinguals have “mental confusion”, especially in rural areas.

Since 1923, much research has been planned, carried out and communicated about the behavioral, cognitive and neurobiological characteristics of bilingualism. Today the idea that bilinguals are confused and are two monolinguals in one person was left behind. The work by Grosjean was very influential in modifying this point of view. Grosjean (1988, p. 234) used to define bilinguals as “those who use two languages in their everyday lives, move in and out of different speech modes depending on the interlocutor they are facing and the situation they are in”. More recently, he included dialects in his definition (Grosjean, 2008, p. 10). He still agrees with this definition in more recent works (Grosjean, 2022, p. 11) and explains “bilingual” may characterize an individual while “multilingual” may describe a country or society. Other researchers agree that multilingualism is “‘the use of more than one language’ or ‘competence in more than one language’” (Clyne, 2017, p. 301).

Kroll *et al.* (2015, p. 1-2) presented briefly a more general definition: “we consider anyone who actively uses two languages to be bilingual, but we also acknowledge that not all bilinguals are the same”. This excerpt is important to mention that, since around 2015, researchers have been rethinking the concepts of bilingualism and depicting it as a continuum of bilingual experiences instead of a category (Poarch; Bialystok, 2015).

In this study, I adopt this definition, which was adapted from the ones mentioned above: A bilingual person has sufficient knowledge of two languages or dialects in order to use them in their lives; this language use may require that only one of the languages be employed according to specific social contexts, activities or groups of people.

2.2.1 Bilingual and multilingual language processing

Language processing is inevitably different for monolinguals and bilinguals. Actually, “the nature of bilingual representation is the result of a highly dynamic and competitive process in which early learning constrains later development, therefore shaping the time course and structure of later language systems” (Li; Zhao, 2013, p. 13). These constraints do not refer to obstacles in learning a language but to the processes recruited and the way both language systems are organized. These dynamics can be observed in both children and adults becoming bilingual.

According to Byers-Heinlein and Lew-Williams (2018), from birth, monolingual and bilingual children prefer linguistic over non-linguistic input and are aware that language conveys meanings. Bilingual infants gradually develop two independent systems for communication and are able to differentiate between languages. During their first year, children become less sensitive to sounds from other languages and more sensitive to sounds from their native language(s) (Kuhl *et al*., 2007). Also, this sensitivity may be influenced by the distance across languages. By the first year, children can perceive words based on the sound combinations, the speakers and the languages they are familiar with (Byers-Heinlein; Lew-Williams, 2018). They use statistical learning to extract patterns of sounds and their frequency of occurrence with help from language-specific cues and social interaction (Romberg; Saffran, 2010). Infants learn words using examples, associations, gestures, eye-gaze, pointing, presence in the room, and attentional cues, and they expect new words to refer to more general or new things. Monolinguals expect that an object has only one label; however, this is not the case for bilingual infants, who learn two or more labels for one object (Byers-Heinlein; Lew-Williams, 2018). Children develop receptive and productive vocabularies. Monolinguals know around

100 words by the age of 12 months and 550 by 18 months (Mayor; Plunkett, 2011). Bilinguals tend to know fewer words since they have vocabularies for both languages, but the total number of words is comparable with monolinguals — but see De Houwer, Bornstein, and Putnick (2014) and Serratrice (2019).

Moreover, children understand words more easily when they are embedded in sentences (Byers-Heinlein; Lew-Williams, 2018). This speech processing ability predicts later language development and is affected by language exposure at home. In their first year, children recognize mispronunciations of words. Bilingual infants seem to ignore mispronunciations when the languages or dialects involved present many variant pronunciations. Children perceive different accents and prefer their community's accents over their parents'. It is relevant to mention these developmental differences between monolinguals and bilinguals because they are related to distinctions in results of research experiments involving these groups.

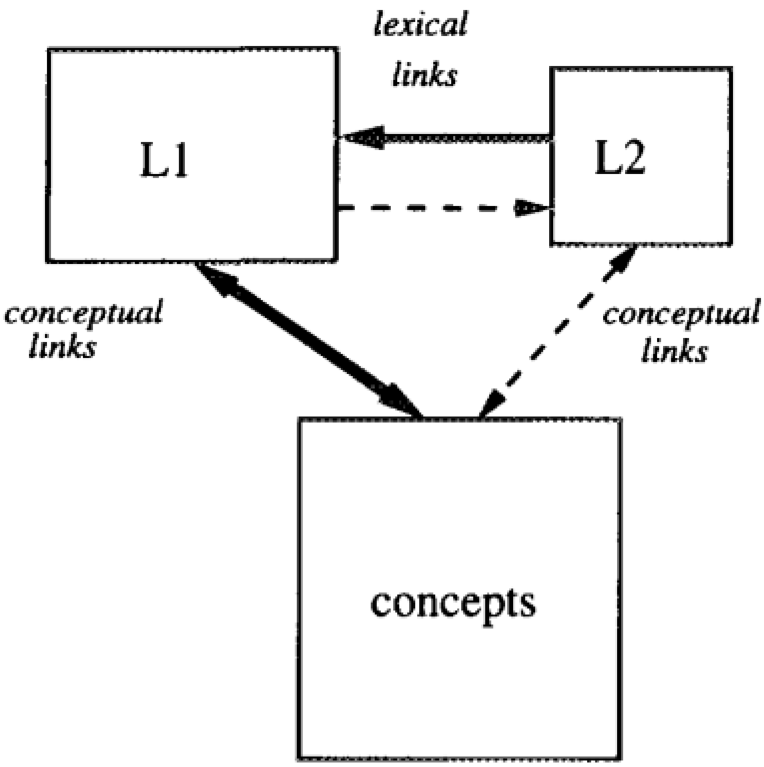
Visual and spoken word recognition for monolinguals are influenced by many input characteristics. These same features also have an impact in these processes for bilinguals. However, in the case of bilingualism, the complex interaction among these characteristics is multiplied due to the existence of more and different linguistic information. As Jared (2015) and Kroll and Ma (2018) mention in their chapters, there is enough evidence to show that the languages of a bilingual person are activated in parallel at least partially.

Over the last 30 years (Jared, 2015), specific types of words have been shown to have different outcomes in spoken and written bilingual processing. The direction of these effects relies heavily on task demands, as emphasized by Dijkstra (2005, p. 191). For example, in lexical decision tasks and naming tasks, cognate words usually produce facilitation effects (Dijkstra; Van Jaarsveld; Ten Brinke, 1998; Peeters; Dijkstra; Grainger, 2013; Schwartz; Kroll; Diaz, 2007;); homographs commonly lead to inhibition effects (De Groot; Delmaar; Lupker, 2000; Dijkstra; Van Jaarsveld; Ten Brinke, 1998); homophones have been associated with both facilitation and inhibition effects (Haigh; Jared, 2007; Lemhöfer; Dijkstra, 2004), and words with dense orthographic neighborhoods usually produce inhibition (Grainger; Dijkstra, 1992; Midgley; Holcomb; van Heuven; Grainger, 2008).

To some extent, these outcomes were replicated in sentence context. However, contextual and syntactic factors play a role in this case. For cognate words, the facilitation effect is still present, but diminished due to sentence constraints such as contextual information (Libben; Titone, 2009). Yet, since task demands impact the direction of the effect, cognates can also have inhibition effects in sentences. Arêas da Luz Fontes, Yeh and Schwartz (2010) observed that, in sentence context, bilinguals had more difficulty in rejecting the irrelevant meaning of an ambiguous word when this word was cognate, that is, shared across languages. All these complex interactions have been informing the creation of bilingual language models.

The Revised Hierarchical Model (RHM; Kroll; Stewart, 1994) was created as a model of transfer (Kroll; Ma, 2018) to explain asymmetries in bilingual translation and may now be considered a model of word production (Kroll *et al.*, 2010). The model proposed that the L1 would mediate access to concepts from the L2 — that is, when using their L2, a bilingual person would need to rely on their L1 in order to access meanings of words (see Figure 4). In the original model from 1994, there was no assumption of lexical non-selectivity or parallel activation of languages, which are well accepted today (Kroll *et al.*, 2010). However, L1 translations would be active when translating from L2. The RHM explained that low-proficiency bilinguals would rely more on the L1 than high-proficiency bilinguals, who would be able to access concepts via the L2 without mediation (Kroll; Ma, 2018, p. 297). In other words, low-proficiency bilinguals would be more dependent on L1 translations while high-proficiency ones would be able to process meaning directly. Indeed, it was seen that low-proficiency bilinguals were more sensitive to word forms than to meaning while high-proficiency ones were more perceptive of semantic aspects than form (Kroll; Ma, 2018, p. 298).

Figure 4 – The Revised Hierarchical Model



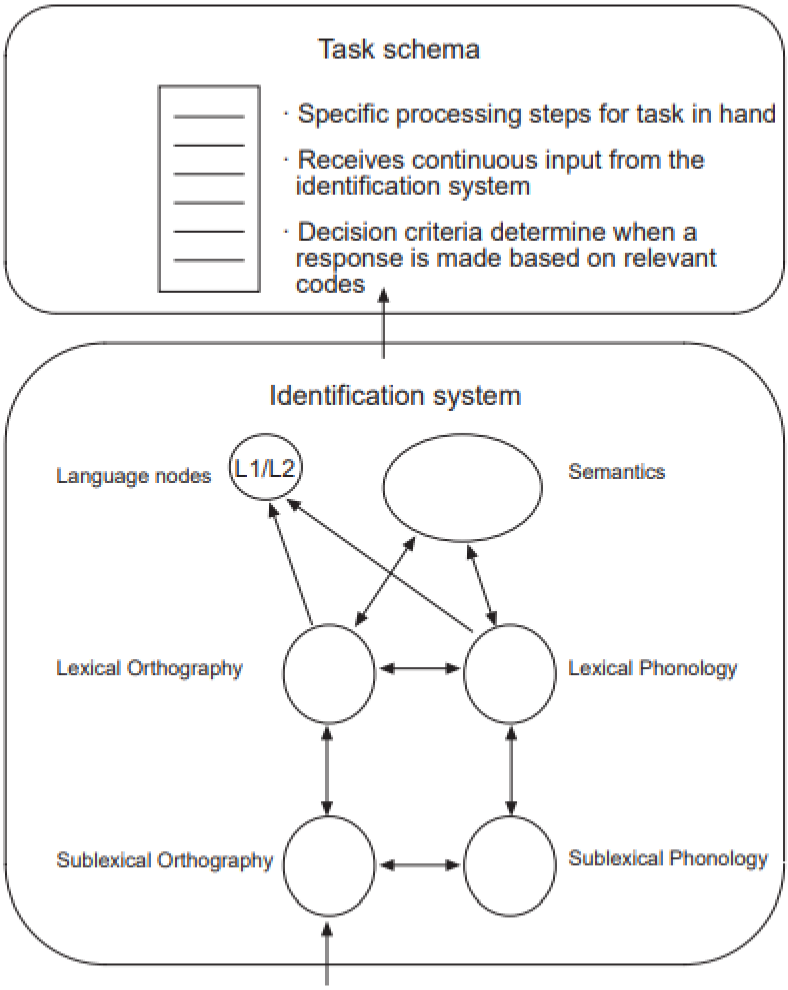
Source: Kroll and Stewart (1994, p. 158)

The model also intended to explain asymmetries in forward and backward translation. That is, in order to translate a word from L1 to L2 (forward translation), the conceptual level would necessarily be accessed. However, when translating a word from L2 to L1 (backward translation), lexical information would be enough and semantic information would not be needed. However, "the RHM's assumption of L1 translation mediation for comprehending the meaning of the L2 word was incorrect" (Kroll *et al.*, 2010) and may be actually seen during language production.

The RHM has received many critiques over the years. A few pieces of evidence against the predictions of the model are the following. For example, Kroll and Ma (2018, p. 298-299) show that low-proficiency bilinguals do indeed access the meaning of L2 words without L1 mediation and that high-proficiency bilinguals not only do access L1 translations but also seem to do it even after accessing the meaning of the L2 word. These results mean that the dynamics of the bilingual lexicon are much more complex and that "the L1 translation is active for all L2 speakers under a range of circumstances" (Kroll; Ma, 2018, p. 299). In addition, although the model assumes that underlying semantic representations are shared across L1 and L2, these links may be influenced by language use, which would lead to different types of organization of these connections (Kroll *et al.*, 2010). Thus, despite the original RHM not predicting cross-language activation, this could be added as an extension to the model.

Diverging from the RHM, there are models which assume language co-activation from the beginning. The Bilingual Interactive Activation Plus model (BIA+, Dijkstra; van Heuven, 2002) is an extension of the Bilingual Interactive Activation model (BIA, Dijkstra; van Heuven, 1998), which was in turn inspired by the IA model presented in section 2.2.2. The BIA+ is a connectionist and localist model for visual word recognition. It is organized in two parts: an identification system and a task schema. The identification system is based on the previous proposal from the BIA model, adding an adaptation of the language nodes. The task schema is directly related to the Inhibitory Control model by Green (1998).

Figure 5 – The Bilingual Interactive Activation Plus model



Source: Dijkstra and van Heuven (2002, p. 182)

BIA+ is data-driven and mainly influenced by bottom-up processes. When the model is presented with a letter string, sublexical information is detected and, in turn, sends activation to word candidates which match with the input. Lexical candidates from both languages are activated not only according to their similarity to the input but also relative to their resting-level activation, which is influenced by frequency, proficiency, use etc (Dijkstra; van Heuven, 2002, p. 182). Candidates spread activation to other levels of the models, such as phonological and semantic. This means the model predicts that, when reading, the activation of phonological and semantic information happens later than of orthographic information — the temporal delay assumption.

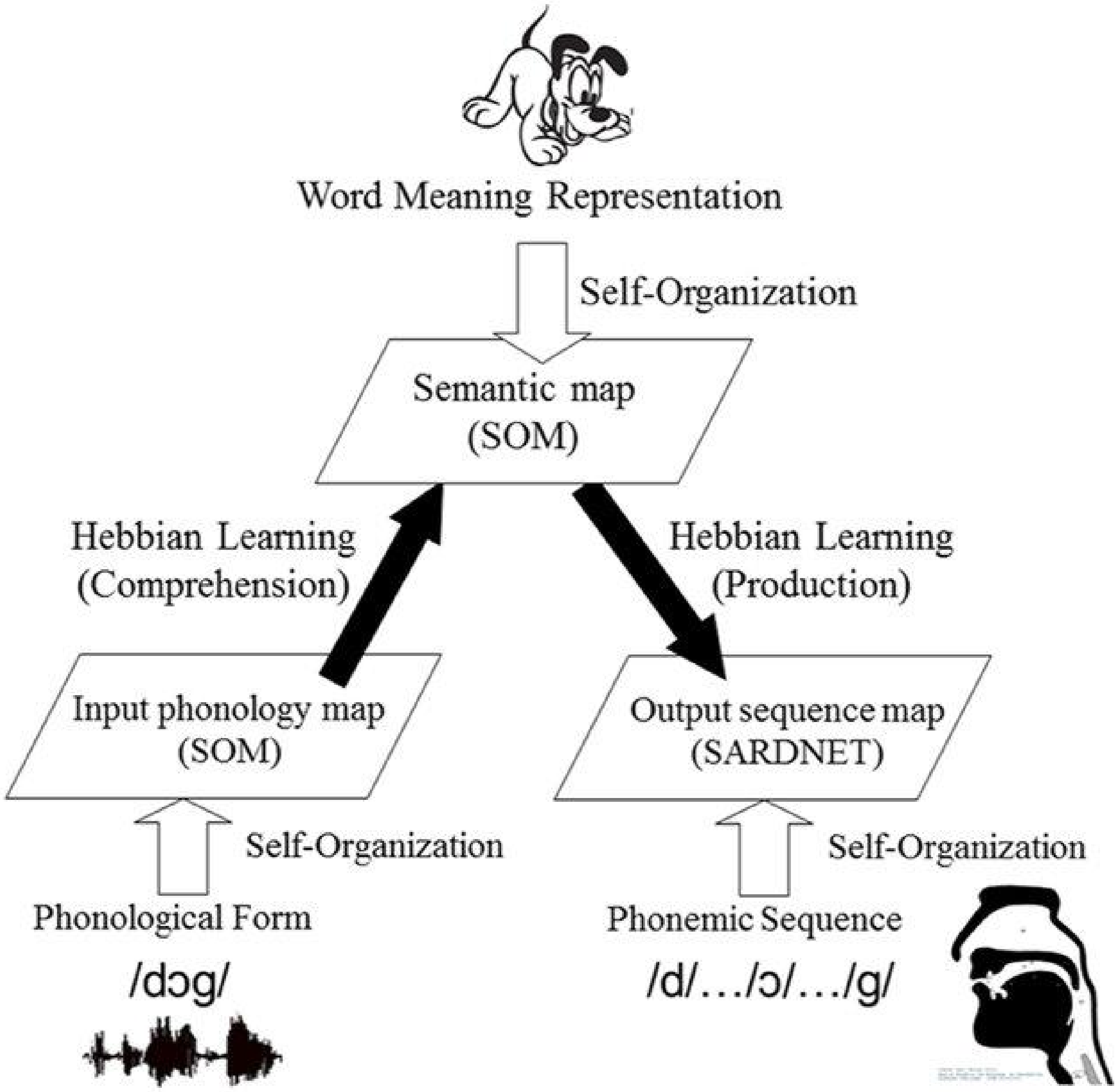
Words will feed activation to the language nodes, which will deal with the global activation of representations from one language or the other. Thus, determining the membership of the word to one language will take place later in the word recognition process. Linguistic context effects, such as sentence context, can only affect the identification system, while non-linguistic effects, such as participants' expectations or task demands, can only impact the task schema. The identification system is also interactive with the parser, which means that word recognition may be influenced by syntactic information from both languages.

Language control and non-linguistic effects would take place later in word recognition, with decision making in the task schema. However, bottom-up effects have been shown to be much stronger than top-down inhibition effects in word recognition, that is, it seems to be almost impossible to suppress activation from the non-target language. So far, BIA+ predicts that bilingual word recognition is mostly unaffected by context or expectations. The task schema is updated according to task demands and participants' strategies and sets the steps which need to be followed. These steps are usually adjusted during the task. Then, when criteria are met, the task schema triggers a response/decision.

BIA+ assumes that lexical access is non-selective and that the lexicon is integrated across the languages of a bilingual person. This means word recognition is affected by degrees of similarity of orthography, phonology and semantics (Dijkstra; van Heuven, 2002). Possible word candidates from both languages are activated in parallel and compete for selection. The temporal delay may lead to stronger cross-language effects from L1 to L2 than from L2 to L1 and to the absence of phonological or semantic cross-language effects (Dijkstra; van Heuven, 2002, p. 183-184). BIA+ seems to have "a special type of representation for cognates, possibly with a strong feedback connection from semantics to orthography" (Dijkstra; van Heuven, 2002, p. 184) which might include morphology (Dijkstra; van Heuven, 2002, p. 185). It is plausible to think that cognates have connections across languages because evidence from fMRA studies hint to it: "word meanings in the two languages converge on the same neuronal populations" (Kroll; Ma, 2018, p. 305). In addition, BIA+ predicts that effects from homographs and cognates arise from interactions within the mental lexicon before the response is produced. Finally, BIA+ may be generalized to spoken word recognition as well since it shares mechanisms with spoken word recognition models and since the auditory processing also seems to be language non-selective (Thierry; Wu, 2007).

While most of the hypotheses in this study are guided by BIA+, another model might offer complementary insight. The Developmental Lexicon II (DevLexII) (Li; Zhao, 2013; Li; Zhao; MacWhinney, 2007) is a connectionist model for spoken word comprehension and production and for language acquisition and processing in both monolingual and bilingual contexts (Li; Zhao, 2013, p. 5). It is composed of multiple self-organizing maps (SOM) connected via Hebbian learning, which very briefly is the hypothesis that the activation of one neuron by another is more efficient when they have repeatedly been activated together. There are three representation levels in the model: a phonological form level for the input, a word meaning level, and a phonemic sequence level for the output.

Figure 6 – The DevLexII model



Source: Li and Zhao (2013, p. 5)

The SOM in the model are an unsupervised learning algorithm which maps the patterns of the input on their two-dimensional structure. Units and their neighboring units are activated for a given input until they resemble this input. Then, this same pattern of units will respond more strongly to similar inputs in the future. Thus, a region of units gets activated in the SOM when receiving input, and similar inputs will be mapped to nearby regions. The activation of a word's phonological form in the first SOM activates its meaning in the semantic SOM, which may trigger its sequence of phonemes for production (Li; Zhao, 2013, p. 5-6). The mapping of the inputs starts off with random and diffuse patterns which will gradually adjust to more focused patterns according to the input's characteristics. DevLexII is able to account for "vocabulary spurt, aspect acquisition, AoA effects, and cross-language semantic priming" (Li; Zhao, 2013, p. 12).

One of the products from DevLexII is the topographical map, which represents the input in regions according to similarities in their patterns of activation and weights. An enlightening test would be to verify how cognate words are mapped onto the 2D structure. What properties have more influence in this organization? This might guide studies into disentangling whether cognate words are special due to morphology (Sánchez-Casas; García-Albea, 2005) or due to a simpler meaning-and-form overlap (Dijkstra *et al.*, 2010). I will describe the proposal that cognate words are a special kind of morphological relationship and present the language co-activation account, which is key for this study, in the next section.

2.2.2 Language coactivation and cognate words

"[I]n bilingual memory the lexical representations of translation equivalents, also those of noncognates, are connected [via meaning]" (de Groot; Nas 1991, p. 109). One specific type of translation equivalent is cognate words. In Psycholinguistic studies, cognates can be considered "words that have the same or similar spellings and/or pronunciations in both languages and have the same meaning" (van Assche *et al.*, 2020, p. 43). They have been studied since 1969 in order to better understand the organization of the bilingual mental lexicon. Cognate effects are seen in reading tasks and in picture naming as well, which indicates that the effects may come from the overlap of orthography and/or phonology and meaning across languages (Kroll; Ma, 2018). In one point of view, cognate words may be treated as special due to "a high level of correspondence between form and meaning across the bilingual's two languages" (Kroll; Ma, 2018, p. 304). In another perspective, cognates are special because they are morphologically related across languages, that is, they share a common root (Sánchez-Casas; García-Albea, 2005, p.227). Both angles agree that cognates are an indication that there is at least some integration across the languages of a bilingual person.

The set of 3 experiments by Preston and Lambert (1969) can be considered the first work which showed evidence of language co-activation in bilinguals. They do not use the term "co-activation", but they do talk about interference of one language onto the other. They used 3 versions of a bilingual Stroop color-word task to answer the following question: "Does the activation of one language system make the other language system inoperative?" (Preston; Lambert, 1969, p. 295).

In experiment 1, 8 balanced English-French and 8 balanced English-Hungarian bilinguals performed the bilingual Stroop task with 6 different sets of cards: in the first set, participants had to say the asterisks color in English; in the second set, participants had to say the color of English color-words in English; in the third set, participants had to say the color of the L2 color-word in English; in the fourth set, participants had to say the color of English color-words in the L2; in the fifth set, participants had to say the color of the L2 color-words in the L2, and in the sixth set, participants had to say the asterisks color in the L2. Since this was a Stroop task, responses mostly showed interference. However, less interference was observed when English-Hungarian bilinguals answered in Hungarian to English words. Preston and Lambert explained that Hungarian words were less similar to English words than French ones. This dissimilarity could be the reason for less interference to occur. So, the degree of similarity was tested in the next experiment.

Experiment 2 has 16 balanced English-German bilinguals. Half of the participants were presented with version 1 of the 6 sets of cards, which had words similar in form between languages, and the other half was presented with version 2, which had words which were less similar between languages than the ones in version 1. The sets of cards had the same organization as in experiment 1, only involving English and German this time. Results showed less interference when words were less similar across languages than when they were more similar, and when stimuli language and response language were the same than when they did not match. So, stimulus characteristics seem to play a role in interference.

Preston and Lambert also tested different proficiency levels. In experiment 3, 8 unbalanced English-French bilinguals performed the bilingual Stroop task with 4 new sets of cards: one with English color-words, one with French color-words, one with English non-color words, and the last one with French non-color words. There was also a set with random sequences of letters. Whenever participants answered using the language they were more proficient in, that is, English, there was always less interference than when they used French. Color words also produced more interference than non-color words. Preston and Lambert explain that it seems that participants' "tendency to translate [task stimulus] is greater when the equivalents [in the other language] have similar stimulus characteristics" (Preston; Lambert, 1969, p. 300). They conclude that there are three factors influencing bilingual written language processing during the bilingual Stroop color-word task: efficiency for encoding and for decoding, and stimulus similarity between translation equivalents.

After Preston and Lambert showing that “activation of a set of processes in one language does not make the other language system totally inoperative” (Preston; Lambert, 1969, p. 301), the study by Cristoffanini, Kirsner and Milech (1986) seems to be the first one to test bilingual language processing with cognate words. They carried out 2 experiments, one with a lexical decision task and the other with an incidental memory task. The objective was to investigate whether 4 types of cognate words across languages would have an effect on repetition priming and whether morphology would be a plausible criterion for the mental lexicon organization.

In experiment 1, 18 Spanish-English bilinguals went through 2 experimental phases: first a study phase, when they named lists of Spanish and English words, and second a test phase, when they completed a repetition priming lexical decision task. In the test phase, English words from the study phase were presented again in English, while Spanish words were presented in English, and new English words were added. There were 5 categories of stimuli: identical cognates (e.g., reunion/reunion), 2 types of morphological cognates with regular suffix (e.g., observacion/observation, crueldad/cruelty), morphological cognates with irregular suffix (e.g., itinerario/itinerary), and morphologically unrelated translations (e.g., tristeza/sadness). Results showed repetition priming (facilitation) both within- and between-languages, with a larger effect for within-languages. The facilitation was observed in all types of words (including noncognates) in the within-language condition, but it was only observed in the 4 types of cognate words in the between-language condition. This indicated that morphology does impact lexical decisions, while language alone does not. Then, Cristoffanini, Kirsner and Milech ran experiment 2 to investigate if cognate words would have a disadvantage in the memory task, which would suggest that cognates are organized by morphological characteristics.

In experiment 2, 8 Spanish-English bilinguals completed a study phase and a test phase. First, they studied 96 Spanish and English words, which were taken from experiment 1, including all stimuli types, except for identical cognates. Then, they were presented with 96 words, half of which were being presented in the same language as in the study phase, while the other halt was being presented in the other language. They were instructed to answer whether words were being presented in the “same” language as before or in a “different” one. Results showed clearly that, when words were presented in a different language and they were cognates, performance in the task was significantly lower than when words were in the same languages and were noncognates.

In sum, both experiments by Cristoffanini, Kirsner and Milech (1986) demonstrate that the languages of a bilingual person interact and the lexicons are probably integrated to some extent. In relation to cognate words, if they were not governed by morphological features, then the facilitation seen in experiment 1 and the interference seen in experiment 2 would not have been observed. However, Cristoffanini, Kirsner and Milech's argument regards the more general hypothesis that morphology, rather than language itself, dictates lexical function. In this dissertation, my interest is in the co-activation of languages of a bilingual person, and both experiments reported by Cristoffanini, Kirsner and Milech (1986) point in that direction.

Another good illustration of the parallel activation of the languages of a bilingual person are the two experiments presented by Marian and Spivey (2003), trying to replicate their previous work (Spivey; Marian, 1999).

The paradigm they used to test language co-activation required participants to move one object from a set of 4 objects while having their eye movements tracked. The names of the objects had similar phonological onsets either within- or across-languages. For example, a between-language experimental word pair was *marka* and *marker*: if languages were co-activated, when they were instructed to move the *marka*, which means *seal* in Russian, the competitor item *marker* in English might be activated. There were also within-language experimental word pairs, such as *speaker* and *spear* and *spichki* and *spitsy*. In a trial, there were 4 items: the target object/word, the competitor one, and two fillers. Location of the object on the table, word frequency, phonological overlap, and physical characteristics of the objects were controlled for.

In experiment 1, 15 Russian-English balanced late bilinguals participated. They performed the task both in Russian and in English, and they received instructions in each language separately in order to increase their activation. Eye movement proportions showed that the participants looked more at the competitor item than to the control fillers within both languages and also across them. These results showed that not only phonological information for a competitor word from the same language as the task was being accessed, phonological information for a competitor word from the other language, which was not being required for the task, was also being accessed. In experiment 2, 12 monolingual English speakers performed the same task in English as the bilinguals. Results showed that monolinguals looked at the within-language competitor item significantly more than at the within-language control filler.

These experiments demonstrate that both monolinguals and bilinguals experienced competition from English words which had some phonological overlap. However, only bilinguals displayed competition effects from English and Russian words which presented phonological overlap. In sum, Marian and Spivey (2003) provide robust evidence for the hypothesis that the languages of a bilingual person are active in parallel during spoken language processing.

De Groot and Nas (1991) were interested in the repetition priming effect between languages and they used cognate words in this paradigm. Participants were Dutch-English simultaneous bilinguals, which were called compound bilinguals at the time, who took part in 4 experiments: experiments 1 (unmasked) and 2 (masked) had only cognate words as stimuli, and experiments 3 (unmasked and masked) and 4 (masked) had both cognate and noncognate words.

The materials for experiments 1 and 2 were the following. Four word-pair lists for primed lexical decision tasks, with 168 prime-target pairs in each list, among which 84 were word pairs and 84 were pseudoword pairs. There were 3 types of primes: repetition, association, and unrelated. One third of the word pairs had their prime selected from association norms; one third had their prime unrelated to the target; and one third had their prime identical (repetition) to the target. Targets were selected from sets of words with similar RTs means and error rates in a baseline unprimed lexical decision task in each language. In total, there were 4 stimuli lists: one list with both primes and targets in English (within-language); another with both primes and targets in Dutch (within-language); another with primes in Dutch and targets in English (between-language), and the final one with primes in English and targets in Dutch (between-language).

Experiment 1 had a total of 72 participants, 18 for each list of prime-target pairs. Results showed that responses to Dutch (the L1) targets were significantly faster than responses to English targets. In addition, responses to repetition primes were faster than to association (related) primes, which in turn were faster than to unrelated primes. The effect of repetition primes was bigger in the within-language lists than in the between-language lists.

Experiment 2 also had a total of 72 participants, 18 for each list of prime-target pairs. This time, a mask of hashes was presented before the primes (a forward mask). All the other aspects were the same. results were similar but smaller than in experiment 1. Responses were significantly faster in the within-language list than in the between-language list. In addition, responses to repetition primes were faster than to association primes, which in turn were faster than to unrelated primes. Again, the effect of repetition primes was larger in the within-language lists than in the between-language lists.

De Groot and Nas (1991, p. 102) explain that “the effect [of priming between languages] has to be attributed to spreading activation in the lexical representational structure of the bilingual, thus revealing aspects of this structure”. But since the stimuli lists which produced these effects contained only cognate words, they modified the materials for experiments 3 and 4 and also added noncognates. This time, there were 2 lists of 180 word-pairs each, 90 words and 90 pseudowords. Targets were the same in both lists and were selected from sets of words with similar RTs means and error rates in a baseline unprimed lexical decision task in English. One list had both primes and targets in English (within-languages), and the other had Dutch primes and English targets (between-languages). There were 3 types of primes once more: repetition, association, and unrelated. One third of the word pairs had their prime selected from association norms; one third had their prime unrelated to the target; and one third had their prime identical to the target. Cognates and noncognates were also rated for similarity.

Experiment 3 had, in total, 68 participants, 17 assigned to each condition: list (either within-language or between-language) and prime (masked or unmasked). Results showed that responses were significantly faster within languages than between languages; responses to cognate words were faster than to noncognates; and responses to repetition primes were faster than to association primes, which were in turn faster than to unrelated. The repetition and the association effects were smaller in between than within languages. When words were unmasked, this repetition effect was also smaller for cognates than noncognates in the between-languages list; when masked, the repetition effect between languages was larger for cognates than noncognates. The effect of masking on cognates was smaller than the effect of masking on noncognates.

Experiment 4 tried to replicate the repetition priming effects in the masked condition. There were 76 participants, 19 in each of four conditions: either all English words list or Dutch-English words list and either primes in uppercase and targets in lower case or primes in lower case and targets in upper case. Results showed that there was no relevant effect of upper or lower case. Responses to within-language targets were slower than to between-language targets. Responses to cognates were faster than responses to noncognates. Once more, it was faster to answer to repetition trials than to association trials, and faster to answer to association trials than to unrelated trials. Experiment 4 replicated the between-languages repetition priming for both cognates and noncognates. However, the association priming between-languages disappeared in this experiment for noncognates, which may mean that "a between-language associative [related] priming effect is restricted to cognates" (De Groot; Nas, 1991, p. 112).

De Groot and Nas propose that cognate and noncognate words are organized differently in the bilingual mental lexicon. They suggest that cognate words share representations at the conceptual level across languages while noncognates have separate representations at this level, one for each language. However, this separation is not sharp: "these data provide a clear support for an integrated bilingual lexical memory" (De Groot; Nas, 1991, p. 113). They also suggest that the repetition priming may be caused by episodic traces, especially in the unmasked conditions, from orthographic similarities. The fact that association priming was reliably seen only for cognate words between languages in experiment 4 reinforces the idea that cognate words share conceptual representations between languages. Moreover, when De Groot and Nas considered that the repetition priming effect was larger than the association priming effect, they concluded that this is a plausible consequence from cognates sharing conceptual representations. In parallel with the RHM, a bilingual person would be expected to access the orthographic information of the cognate words first, then the conceptual node, and then the word in the other language, that is, 3 steps would be needed. However, for the repetition priming effect, since the same orthographic form is presented twice, there would be no need to access the conceptual node, and the previous activation of the word form would suffice. This evidence — and others (Cristoffanini; Kirsner; Milech, 1986) — will lead to the hypothesis that cognate words share morphology across languages.

In their chapter, Rosa Sánchez-Casas and José García-Albea (2005) presented a comprehensive review of studies investigating the bilingual lexicon via cognate words. Their main argument is that the mental lexicon is organized according to morphological characteristics and so is the bilingual one. Consequently, cognate words can be considered a special type of this morphological relationship since the effects they seem to cause are not entirely explained by overlap in form and meaning. The body of research they summarize in this chapter point to two approaches towards the representation of cognate words: a basis in morphology, which includes meaning, or a basis in form (orthographic and/or phonological) and meaning.

The first bilingual studies with cognate words used priming paradigms, in which the prime word belonged to one language and the target word belonged to the other one. Most studies were interested in repetition priming — when primes are identical to targets — and in associative priming — when primes are semantically related to targets. There were studies which showed that cognate words facilitate answers, but there were also studies which demonstrated that noncognates can lead to facilitation. For example, de Groot and Nas (1991), previously presented here, observed facilitatory repetition priming effects from both cognate and noncognate words, but associative priming effects only for cognates, not for noncognates. Williams (1994) and Sánchez-Casas and Almagro (1999) also saw facilitatory effects for noncognates, while at least five other studies did not see this impact for noncognates. The facilitation produced by noncognate translation words was unexpected since they do not share form across languages.

Cristoffanini, Kirsner and Milech (1986) (as previously described here), Davis, Sánchez-Casas and García-Albea (1991), García-Albea, Sánchez-Casas and Valero (1996), García-Albea, Sánchez-Casas and Igoa (1998), and Sánchez-Casas, García-Albea and Igoa (2000) showed that across languages cognate words had facilitation effects and that noncognates had no effect. Most surprisingly, Davis, Sánchez-Casas and García-Albea (1991) observed that the facilitation from cognates was not modulated by orthographic overlap. In addition, the degree of meaning overlap does not seem to influence recognition of cognates and noncognate words (Guasch, 2001; Sánchez-Casas; Suárez-Buratti; Igoa, 1992). Most studies including false friends indicate that they cause interference/inhibition effects, not facilitation. However, at least Gerard and Scarborough (1989) and Sánchez-Casas and Almagro (1999) observed facilitation for false friends in a lexical decision task.

Taking all these results together, Sánchez-Casas and García-Albea (2005, p. 235) explain that neither form overlap nor meaning overlap are enough separately to explain facilitation effects of cognate words in cross-language priming. Both have an impact in lexical access: “Form seems to play a role early on in the process [...] In contrast, meaning similarity by itself seems to exert an influence later in the recognition process”. Thus, it is plausible to consider cross-language cognate words as morphologically related.

The study by Costa, Caramazza, Sebastian-Galles (2000) is one of the first to use cognate words in order to deliberately avoid using words from both languages in the same experiment. That is, stimuli would contain words from online one of the bilinguals` languages, which would activate that one language; however, if there is parallel activation of both languages and cognate words share lexical representations, then cognate words would activate the nontarget language.

Costa, Caramazza and Sebastian-Galles (2000) present two experiments using a picture naming task with cognate and noncognate words. In experiment 1, 21 Catalan-Spanish bilinguals and 21 Spanish monolinguals completed a picture naming task in Spanish. Stimuli were either cognate across languages, that is, very similar phonologically, such as *gato* in Spanish and *gat* in Catalan, or noncognate, that is, phonologically dissimilar, such as *mesa* in Spanish and *taula* in Catalan. Results showed that only for bilinguals the cognate pictures were responded to faster than the noncognate ones, that is, cognates produced facilitation, which is in accordance with cascade activation models. More specifically, phonological segments of words that are shared across languages are activated together, which activates both words.

In experiment 2, the same task was completed by 23 Spanish-Catalan and 23 Catalan-Spanish bilinguals, who were either dominant in Spanish or in Catalan. Results were similar to experiment 1, that is, a cognate facilitation effect was seen during picture naming. This time, the facilitation was larger for bilinguals who were naming in their nondominant language. Once more, Costa, Caramazza and Sebastian-Galles (2000) explain that these results are accounted for by cascaded activation models and also by the RHM by Kroll and Stweart (1994).

Schwartz, Kroll, and Diaz (2007) had 18 English-Spanish bilinguals name (read out loud) English and Spanish words. Half of them were cognate across languages, the other half were not. However, cognate words were selected according to the degree of orthographic and phonological overlap: for example, more orthographic and more phonological overlap (+O +P), or less orthographic, more phonological overlap (-O +P). Results showed that naming latencies in Spanish were as follows: from the faster to the slower, +O +P < +O -P = -O -P < -O +P. These differences in latencies indicate that responses were influenced by the degree of overlap both of orthographic form and of phonological form. Thus, the co-activation of cognate words across languages happens via the nuanced interaction of many input characteristics.

Many studies investigated the extent to which each word property would impact co-activation. Lemhöfer *et al.*, (2008) examined how within- and between-language variables affected the influence of the L1 on an L2 progressive demasking task. Within-language variables were word frequency and length, morphological family size and orthographic neighborhood, bigram frequency, number of meanings, syntactic ambiguity, concreteness, familiarity, and meaningfulness. Between-language variables were orthographic neighborhood, and interlingual cognates and homographs. Participants were French-English, German-English and Dutch-English bilinguals who completed the task in their L2 in three separate sessions. In general, correlations and regression results showed that L2 word recognition seems to be mostly influenced by within-language factors. However, there are differences when comparing L1 and L2 language processing. Bilinguals were affected by both spoken and (specially) written word frequency while native speakers of English were influenced only by spoken word frequency. Facilitation from morphological family size was larger than from word frequency for bilinguals but not significant for natives. Both bilinguals and natives were slowed down by a bigger number of higher frequency neighbors, and were helped by word meaningfulness and familiarity. Syntactic ambiguity helped bilinguals, but not natives. Interlingual neighborhood size and homographs seemed to have no effect on bilinguals, but cognate words did facilitate recognition. Thus, types of word frequency affect natives and nonnatives differently, together with morphological family size and syntactic ambiguity. And between languages, cognate status was the strongest predictor of L2 word recognition.

Adding to the literature of bilingual reading, Dijkstra *et al.* (2010) intended to detail the impact of cognate similarity across languages via three different tasks. Materials were selected in a rating study when words were classified according to orthographic, phonological and semantic similarity between Dutch and English. In experiment 1, Dutch-English bilinguals completed a lexical decision task. Results showed that reaction times to cognates were facilitated as orthographic similarity and frequency increased, and that phonological similarity facilitated answers to identical cognates. In experiment 2, participants completed a language decision task, and orthographic overlap of cognates produced inhibition, which increased abruptly in the case of identical cognates. Also, English word frequency and phonological similarity did not play a role, but the Dutch ones did. Finally, in experiment 3, participants completed a progressive demasking task, and no orthographic similarity effect was seen. Answers to identical and non-identical cognates were predicted by English frequency and semantic similarity. The exception was low frequency identical cognates, which showed a facilitation effect. The authors concluded that a localist connectionist account, which predicts separate lexical representations for cognate words and influences from orthographic and semantic similarities and frequency, must be the best model to explain these patterns of findings.

Most studies on bilingual reading focus on word and sentence reading. In a distinct perspective, Cop *et al.* (2016) tested the cognate facilitation effect during the reading of an entire novel. Participants read a story, which contained identical and non-identical cognates, half in their L1 (Dutch) and half in their L2 (English), and their eye movements were recorded. Generally, there was facilitation from cognates in both languages. Orthographic overlap across languages had an impact on L2 first fixation duration and skipping rates, especially for short words. Also, identical cognates affected L2 reading times and go past times. On the other hand, in L1 reading, the cognate facilitation effect was smaller; orthographic overlap influenced first fixation durations, and identical cognates had an impact on total reading times, especially for high frequency nouns. These results seem to point to two separate lexical representations for identical cognates across languages. Also, they suggest that higher language proficiency decreases the effect of word frequency.

There are inconsistencies in the literature on the effects caused by cognate words and the co-activation of languages. Focusing this aspect, Lemhöfer, Huestegge, and Mulder (2018) investigated L2 influences on L1 reading through near-cognate words. In Experiment 1, they prepared a lexical decision task with L1 (German) words and nonwords which were homophonous and spelled in the L2 (Dutch). Participants were German monolinguals and German-Dutch bilinguals. L2 near-cognate nonwords caused inhibition for bilinguals, and homophones, for monolinguals. Also, vocabulary in Dutch was positively and moderately correlated with the L2 near-cognate effect. However, there was no effect of near-cognates on L1 word processing. In Experiment 2, the authors then embedded the stimuli in German sentences. L2 near-cognates produced a facilitation effect for bilinguals, according to re-fixation and re-reading times; there was no or little facilitation for monolinguals. Also, the bigger the bilinguals' vocabulary in Dutch, the fewer times they re-fixated near-cognate words. Again, there was no effect of near-cognates on L1 sentence reading. The authors explain that these results support a non-selective view to lexical access even in a sentence context, although L2 impact on L1 seems to be weaker.

The case of cognate words is interesting because it sheds light on the mental representation of the bilingual lexicon. Thus, Comesaña *et al.* (2015) examined the effects of identical and non-identical on the direction of the cognate effect with two experiments using lexical decision tasks. In experiment 1, they included both identical and non-identical cognates in the Catalan-Spanish task, while in experiment 2 they added only non-identical ones. Experiment 1 showed that the cognate facilitation effect was caused by the identical cognates in the stimuli list. Further, there were interference effects in test conditions with high phonological overlap. On the other hand, experiment 2 presented a cognate interference effect and suggested that combining high orthographic *and* phonological overlap produced facilitation. The authors propose that whether identical cognates are added to the stimuli list or not affects the cognate effect direction. This may modify the language context and contribute to higher activation of the language irrelevant to the task. Thus, the bilingual mode was more intense than in the task with only non-identical cognates. Simultaneously, the interference effect in experiment 2 may be explained by lateral inhibition due to less overlap of representations.

This effect may also be influenced by other factors. Arêas da Luz Fontes and Schwartz (2015) were interested in investigating cross-linguistic activation of representations and access to less frequent meanings of ambiguous words. In experiment 1, participants completed a lexical decision task composed of cognate or non-cognate homonym words. In experiment 2, these stimuli were inserted in sentences with constrained or neutral contexts. Results indicated that the disambiguation process in L1 and L2 may be explained by the same cognitive mechanism. In addition, better performance in both experiments was negatively associated with the cognate facilitation effect. It is suggested that some participants relied on word form to complete the task, which led to higher error rates due to the presence of homonym words. The authors theorize that the cognate facilitation effect is modulated by proficiency, in that participants with intermediate proficiency are benefited more than advanced participants.

Moreover, the cognate effect was tested in sentence context. Bultena, Dijkstra and van Hell (2014) were interested in the co-activation of languages during reading. Their objective was to test if the cognate effect in sentence context is influenced by L2 proficiency, word class and task demands. In experiment 1, 37 adult Dutch-English bilinguals read 53 English sentences while their eye movements were being tracked. Most sentences (29) had either a cognate or noncognate noun, and the rest (24) had a cognate or noncognate verb. There were also 75 filler sentences. Participants' L2 proficiency was self-rated. Results showed that nouns were processed faster than verbs and that cognate words facilitated the recognition of nouns for all eye movement measures. However, cognate verbs did not have the same advantage: there was facilitation only for first-fixation duration and a very small effect. In addition, readers with lower L2 proficiency displayed more cognate facilitation than higher proficiency ones.

In experiment 2, Bultena, Dijkstra and van Hell (2014) used the same materials as in experiment 1 to test task demands. This time, it was a self-paced reading task, and 38 adult Dutch-English bilinguals participated. Results once more indicated that nouns were responded to faster than verbs. "[T]he proficiency-dependent cognate effect was present for both nouns and verbs" (Bultena; Dijkstra; van Hell, 2014, p. 1231). The facilitation was larger for nouns than for verbs. Accordingly, the cognate facilitation effects are present in the sentence context and are affected by word class and L2 proficiency. It should be highlighted that the task demands associated with the self-paced reading technique allowed the effect to accumulate. This is one of the reasons for opting for self-paced tasks in study 2 in this dissertation.

In this manner, Allen, Conklin and Miwa (2021) carried out an experiment similar to the one in study 2 here. They examined whether, in a self-paced reading task, cognate words across Japanese and English would lead to a facilitation effect. Also, a lexical decision task was included. Both tasks were in English, and cognate words overlapped phonologically and semantically across languages. Participants were 24 Japanese-English bilinguals and 24 monolingual native speakers of English. L2 proficiency was measured using a vocabulary test. The self-paced task was similar to the one by Bultena, Dijkstra and van Hell (2014). Results demonstrated that participants responded faster in the self-paced task than the lexical decision task. Higher L2 proficiency led to faster responses as well. Among the various factors which had significant effects in a general model, I will highlight phonological and semantic similarities and the frequency of the Japanese counterpart of the cognate target. In the lexical decision task, both phonological and semantic similarities had a significant facilitation effect, but as a main effect. In the self-paced task, results were more nuanced: semantic similarity had no main effect, but phonological similarity did, only for low-frequency Japanese items. In other words, only in the self-paced task, phonological similarity and L2 proficiency had the largest effects for those English words whose Japanese counterpart had low-frequency. This facilitation was not seen when English words had high-frequency Japanese counterparts.

The works by Bultena, Dijkstra and van Hell (2014) and Allen, Conklin and Miwa (2021) indicate that the cognate facilitation effect may still be seen in a sentence context, although it may be smaller than in a lexical decision task. Moreover, they emphasize the impact that L2 proficiency has in mediating this effect. In sum, they both strengthen the body of evidence towards the idea that bilinguals had their languages activated in parallel.

The next study involves co-activation of syntactic structures instead of isolated words. It is mentioned here because it is the only one with a self-paced *listening* task I could find.

Van Dijk, Dijkstra and Unsworth (2022) investigated the co-activation sentence structures across languages during a self-paced listening task. Participants were monolingual 39 Dutch children and 40 English-Dutch children and 42 German-Dutch children, all simultaneous bilinguals. The syntactic structure in question was the passive voice, which has two possible forms in Dutch, one of which is the passive form in English while the other is the passive form in German. Part of the data trimming process van Dijk, Dijkstra and Unsworth used here was used in this study, in the data analysis exploration phase in study 2. Proficiency and language dominance were measured with a sentence repetition task. Participants also completed a digit span task and two subtests on non-verbal intelligence. The self-paced listening task was administered in all of each participant's languages in two sessions. Results showed that bilinguals were faster than monolinguals and that German-Dutch children were faster than English-Dutch ones. In relation to structure co-activation, only German-Dutch children had slower responses to the critical segment, which even slower the more German-dominant they were. Thus, it is possible to see that the shared passive voice structure was activated during tasks which tapped only one of the languages of the bilingual. In this case, when there was sufficient co-activation, it led to an inhibition/interference effect.

It is possible to observe from the studies presented above, most investigations of the cognate effect were conducted with participants who are mostly speakers of majority languages or who are not in a migration context. Thus, I conducted a brief literature search using the search mechanism at the Periódicos da CAPES database. I used the following keyword sets: cognate children language activation (16 entries); cognate adult language activation (36 entries);

cognate minority language activation (7 entries); cognate language activation (214 entries). Among these, five were not studies testing participants responses to cognate words (essays, rating studies, reviews, translation studies), 18 were not testing the cognate effect per se. Many entries were repeated, and many more were actually from Biology and Chemistry studies. After ignoring these cases, there were 72 studies with adult participants about languages other than minority ones, nine studies with children and teenagers as participants about languages other than minority ones, and seven studies about minority or heritage languages specifically.

Among the nine studies investigating cross-language activation via cognate words with children, seven of them showed cognate facilitation effects and only two showed no effects at all. Among the 7 studies investigating cross-language activation via cognate words whose participants were speakers of minority languages, four studies showed cognate facilitation effects, one showed inhibition effects, one showed no effect at all, and one could not be accessed. Only 5 of these studies focused on children and/or teenagers speakers of a minority language. Among these, two studies reported cognate facilitation, one reported inhibition, and one reported no effect. I will be describing the three comprehension studies with children as participants.

Woolpert (2018) was interested in the effects of cognate words and false friends during a word-picture matching task conducted in English. Participants were 72 Spanish-English bilingual children who were heritage speakers and 77 English monolingual ones. They completed decoding, vocabulary and reading comprehension tasks in English (and Spanish for the bilinguals) and the word-picture matching task, which presented a picture surrounded by four words and asked participants to point to the word representing the picture. There were three conditions in the task: a baseline condition, with a target word, an orthographically similar foil and two completely different foils; a cognate condition, with the target word being cognate; and a false friend condition, with one of the foil words being a false friend. Results showed that, for the bilingual children, false friends slowed response times and increased error rates while cognate words seemed to have no impact at all in any behavioral measure. So, for this minority language group of children, cognate words did not influence their performance in a word-picture matching task.

Bosma *et al.* (2019) tested longitudinally how proficiency in the L1 and age would affect how 120 Frisian-Dutch children would respond to a Frisian receptive vocabulary task with cognate and noncognate words. Frisian is a minority language in the Netherlands, while Dutch is a majority language. Participants were divided into three groups according to exposure to Frisian at home. The group with less exposure to Frisian had the highest SES background of the three. Results showed that the low-exposure group performed significantly worse in the task than the other two, which in turn performed similarly. However, the three types of cognate words used in the task lead to significantly higher accuracy scores than noncognates for the low-exposure group. This facilitation effect was modulated by the amount of cross-linguistic overlap — the most overlap, the higher the accuracy scores — and it was still seen 2 years later. The other two groups had similar performances for all word types, which indicates that there was no cognate effect. Bosma *et al.* (2019) do not discuss the fact that the low-exposure group belonged to a higher SES than the middle- and the high-exposure groups, which did not show indications of cognate facilitation.

Bosma and Nota (2020) investigated how cognate words influence sentence reading in Frisian and in Dutch by Frisian-Dutch bilingual children. I could not access the entire article to gather all method and procedures details; however, I was able to read a summary of the results. Bosma and Nota (2020) observed that only identical cognate words caused facilitation during sentence reading and only in Frisian. Non-identical cognates did not produce significantly different reading times than noncognates. This is a result which diverges from what Bosma *et al*. (2019) had seen previously. In their study, in the third school year, low-exposure bilinguals showed similar accuracy scores for identical and non-identical cognates. In Bosma and Nota's (2020) study, facilitation was only visible for identical cognates. This points to the direction that degree of similarity of the cognate word plays an important role in the cognate effect (Dijkstra *et al.*, 2010).

Among the 72 studies investigating cross-language activation via cognate words with adults, 56 studies presented evidence of cognate facilitation, 11 showed cognate inhibition, and six indicated no effects from cognate words. Some studies observed the cognate effect in opposite directions. For example, one study reported both inhibition and null effects, and one showed both inhibition and facilitation effects. One study could not be accessed. Among the 7 studies on cross-language activation via cognate words with minority language speakers, four studies reported cognate facilitation effects, one inhibition effect, one no effect at all, and one could not be accessed for reading. Finally, only two studies invited adult participants speakers of minority languages/dialects and both presented cognate facilitation effects, but only one was a language comprehension study. I will describe this last one here.

Muntendam *et al*. (2022) investigated whether cognate words in relation to noncognates would elicit faster responses in auditory lexical decision tasks in Turkish and in Dutch. They were also interested in the effect of the position of the stressed syllable on responses. Participants were second-generation adult speakers of Turkish and Dutch who reported being more proficient in Dutch than in Turkish although their L1 was Turkish. Results demonstrated that, in the Turkish task, cognate words presented a marginal inhibition effect in relation to noncognates when in both languages the stressed syllable was the penultimate one; in general, cognate words had no significant effect on RTs in comparison with noncognates. In the Dutch task, cognate words had a nonsignificant facilitation effect on the conditions involving words with stress on the last syllable, but had a significant inhibition effect on cognate words stressed on the penultimate syllable, which is less typical for Turkish words. Thus, Muntendam *et al*. (2022) show that heritage adult speakers display cognate inhibition effects when performing an auditory lexical decision task in their dominant L2.

Ultimately, the findings cited in this subsection demonstrate that cognate words are able to indicate language co-activation during word recognition and sentence reading. In addition, it is possible to verify inconsistencies across results with cognates. Some studies revealed a facilitation effect (Arêas da Luz Fontes; Schwartz, 2015; Brenders; van Hell; Dijkstra, 2011; Cop *et al.*, 2016; Lemhöfer *et al.*, 2008). Other works showed divergence in the direction of the effect according to the use of identical and non-identical cognates, some pointing to facilitation, some to inhibition (Comesaña *et al*., 2015; Cop *et al*., 2016; Dijkstra *et al.*, 2010; Lemhöfer; Huestegge; Mulder, 2018). Moreover, studies on the effect of cognate words in sentence context show that there is facilitation, even if it is weaker than with isolated words, and that it is mediated by L2 proficiency levels (Allen; Conklin; Miwa, 2021; Bultena; Dijkstra; van Hell, 2014).

In the next section, I will present a summary of the most impacting events in the history of Haiti, examples of the relationship between Brazil and Haiti, a brief description of the structure of HC as a language, and the most relevant studies about HC for this dissertation.

2.3 HAITIAN CREOLE

2.3.1 Migration from Haiti to Brazil

Until the year of 1804, Haiti was an important French colony. It was called "pearl of the Antilles" because of its numerous plantations, which included coffee, sugar, cotton, indigo, cacao, corn, potato, yam, sorghum, banana, and cassava plantations (Baeninger, 2017). In 1804, it became independent of France and was the first black republic in the world. Later, it went through two dictatorships; the first one was supported by the United States of America in order to maintain Haiti as an ally during the Cold War (de Moraes; de Andrade; and Mattos, 2013).

The military occupation of Haiti with troops from the United States began in 1915 and ended in 1934 (Magalhães; Baeninger, 2017). The influence of this military presence culminated in political and economic impacts. At this point, the UN intervened together with Brazilian troops.

In 1986, the first president of Haiti was elected, Jean-Bertrand Aristide, but he had to flee the country due to a coup d'etat in 1991. The UN and the USA helped Aristide return as president, and he was reelected in 2000. However, because there was suspicion of fraud during the election, riots emerged. Aristide had to be removed from the country, and Bonifácio Alexandre, president of the Supreme Court of Haiti, took the leadership (de Moraes; de Andrade; Mattos, 2013).

In 2004, Brazil participated in the UN Stabilization Mission in Haiti (Missão das Nações Unidas para a estabilização no Haiti, MINUSTAH) together with other 20 countries (CFN, N/Y), which was one of the 9 peacemaking missions Brazil took part in (Lima, 2017). The mission in Haiti was created to help secure political objectives and human rights after the departure of Jean-Bertrand Aristide to exile (CFN, N/Y). The mission also contributed to deal with the outcomes of the earthquake, the destruction caused by hurricane Matthew in 2016, and the outbreak of cholera in 2010; it was ended in October 2017 (UN, N/Y; CFN, N/Y).

This illustrates that cooperation programs between Brazil and Haiti trace back to 2004, when the Basic Agreement for Technical and Scientific Cooperation (Acordo Básico de Cooperação Técnica e Científica) between the two countries was initiated. Since then, 15 cooperation projects have been carried out, including after the desolating seismic shock of 2010 (ABC, 2012). Ten years after the fateful earthquake, Haiti had not managed to overcome the calamitous experience: there were still many political conflicts and an economic crisis (Charles, 2020).

According to the Haitian Government, more than 220,000 people have died in the earthquake of January 12th 2010 (UN, N/Y). An article in the Brazilian Marine website estimates that 250,000 people have died and that 40 thousand were amputated. The earthquake was classified as 7,3 in the Richter scale and was estimated to have compromised around 80% of buildings in Port-au-Prince (de Moraes; de Andrade; Mattos, 2013).

As a comparison, in 2010, Haiti's HDI (Human Development Index) was 0.433, and Brazil's was 0.793. In 2020, Haiti's had increased to 0.540, and Brazil's had decreased to 0.758. More details are presented in Figures 7 to 10 (UNDP, 2023, Human Development Index). In 2021, both HDIs presented some decrease: Haiti's was 0.535, and Brazil's was 0.754 (UNDP, 2023, Human Development Insights).

Figure 7 – 2010 Haiti HDI value

Haiti	
2010 HDI value	0.433
HDI change from 2009	-0.075
Life expectancy at birth	46.0 years
Expected years of schooling	8.7 years
Mean years of schooling	4.7 years
Gross National Income per capita	2,906 (constant 2017 PPP$)

Source: UNDP (2023)

Figure 8 – 2010 Brazil HDI value

Brazil	
2010 HDI value	0.723
HDI change from 2009	+0.006
Life expectancy at birth	73.2 years
Expected years of schooling	13.9 years
Mean years of schooling	6.9 years
Gross National Income per capita	14,414 (constant 2017 PPP$)

Source: UNDP (2023)

Figure 9 – 2020 Haiti HDI value

Haiti	
2020 HDI value	0.540
HDI change from 2019	-0.003
Life expectancy at birth	64.1 years
Expected years of schooling	9.7 years
Mean years of schooling	5.6 years
Gross National Income per capita	2,940 (constant 2017 PPP$)

Source: UNDP (2023)

Figure 10 – 2020 Brazil HDI value

Brazil	
2020 HDI value	0.758
HDI change from 2019	-0.008
Life expectancy at birth	74.0 years
Expected years of schooling	15.6 years
Mean years of schooling	8.1 years
Gross National Income per capita	13,791 (constant 2017 PPP$)

Source: UNDP (2023)

After the political, environmental and social crises mentioned above, the cooperation between Brazil and Haiti was expanded. In 2011, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and other three government institutions created the Emergency Program for Higher Education Pro-Haiti-Undergraduate (CAPES, 2011). The program offered scholarships for students enrolled in higher education institutions in Porto Principe to come to Brazil to complete part of their undergraduate course in a Brazilian institution. The program lasted for 5 years. In addition, in 2012, the Brazil-Haiti Bilateral Program was active and it was composed of three projects related to health demands (ABC, 2012). I could not find information on whether it is still ongoing today or not. There were also three trilateral projects and other two bilateral ones (ABC, 2012). There is also the Sérgio Vieira de Mello Chair, a project involving the UNHCR and 35 universities in Brazil in honor of Sérgio Vieira de Mello, who was a High Commissioner at UN. The chair promotes projects for accessing higher education, research groups and topics relative to refugees, interdisciplinary community services, and taking part in public policies accountability (ACNUR; CSVM, 2002). Moreover, since 2012, Brazil has issued humanitarian visas for Haitian immigrants (Gomes, 2017), "which grants them the right to live, work, study and apply for permanent residency in Brazil" (IOM, 2023, p. 1).

Today "Haiti is [still] facing a devastating humanitarian and safety crisis, marked by systemic violence, acute food insecurity, fuel shortage and limited access to health care and basic sanitation" (UNHCR, 2023). In June 2023, the Brazilian Government and the United Nations High Commissioner for Refugees (UNHCR), which was created in 1950, launched a protection and interaction plan for the Haitian population living in Brazil (UNHCR, 2023). It is estimated that there are about 161,000 Haitian people settled in Brazil, who should have access to rights and to public policies. And in order to secure those rights, the objectives of this new plan are: produce diagnosis about access to rights and opportunities in Brazil; broaden mechanisms of humanitarian reception and of documentation; include socioeconomic integration strategies; and support community structures (UNHCR, 2023).

The exact number of Haitian people living in Brazil is not known because there were many cases of undocumented immigration. For example, Magalhães and Baeninger (2017) interviewed 31 Haitian people living in Balneário Camboriú, Santa Catarina, and only 4 of them reported having entered Brazil with the necessary documents for safely registering as an immigrant. With resolution 97/2012 from CNIg, Brazil established a humanitarian visa for Haitian people who were interested in working and settling in the country. A total of 1,200 visas would be issued annually, 100 per month. However, de Moraes, de Andrade, and Mattos (2013) comment that only 30% of those 100 visas were issued in the first month in Port-au-Prince, because most people did not meet the eligibility criteria even though the demand was high and because there was a fee to be paid. Thus, undocumented entrances seem to be linked to, at least, bureaucratic and social factors.

The estimate of 160,000 Haitian people settled in Brazil is much higher than the official figures. For example, from 2012 to 2016 a total of 73,077 Haitian immigrants registered at Federal Police in Brazil (IMDH, 2016). However, according to Magalhães and Baeninger (2017), in 2017 it was estimated that more than 50 thousand Haitian people were living in Brazil. Moreover, in 2021, 27.88% of migrant workers legally employed in Brazil were from Haiti (Cavalcanti; de Oliveira; Silva, 2022, p. 11).

These numbers can be verified through DataMigra, which is a platform composed of databases with administrative records about international migration and asylum requests. According to SisMigra, a database within DataMigra, since 2010, 173,493 people entered Brazil legally coming from Haiti and got registered. In comparison, since 2010, 2,808 people entered Brazil coming from the Dominican Republic.

SisMigra also presents data since 2011 on the number of people registered in Florianópolis coming from Haiti: 2,457 people. Some participants in this study mentioned that they had stayed in the Dominican Republic for a while before coming to Brazil. Thus, I looked for people entering Brazil from the Dominican Republic as well. SisMigra offers data since 2013 on the number of people registered in Florianópolis coming from the Dominican Republic: 46 people. These numbers are very different from the ones I could find in newspapers: The estimated number of Haitian people living in Florianópolis and surrounding cities in 2015 was around 8,000 people, while in Palhoça, next to Florianópolis, the estimate was 2,000 Haitian people (NSC, 2015). However, SisMigra says only 1,715 people from Haiti have been registered in Palhoça since 2012.

As commented by Gomes (2017), Brazil has had legal support for immigrants since 1997. Regarding the State of Santa Catarina, implementations were made more recently. Magalhães and Baeninger (2017) mention that most immigrants seem to be attracted to the state of Santa Catarina, especially to four cities: Blumenau, Joinville, Chapecó, and Itajaí. In 2016, the Reference and Welcoming Center for Immigrants (Centro de Referência e Acolhimento dos Imigrantes, CRAIS) was created in Santa Catarina with help from Florianópolis City Hall, the State Government of Santa Catarina and the Federal Government (Santa Catarina, 2016). Also, Florianópolis was the second Brazilian state to have specific City Hall policies for immigrants (Guagliano, 2020).

In August 2021 (Kleinebing, 2023), the State of Santa Catarina also launched the Welcoming Program for Migrants and Refugees (Programa de Acolhimento a Migrantes e Refugiados, PARE) with the objective of promoting linguistic integration and developing pedagogical goals of immigrants enrolled in public State schools (CONSED, 2022). The program gathers immigrant students at school during the opposite class shift, that is, if students have regular classes in the morning, PARE classes occur in the afternoon, and vice-versa. The program focuses on language and math classes, with study sessions on the alphabet, orthography, reading fluency, oral text production, comprehension, writing, narrative composition, and math (CONSED, 2022). In the beginning of 2022, there were 6,323 students enrolled in PARE in the entire State (CONSED, 2022); in the second semester of 2023, there were 1,150 students (Kleinebing, 2023).

Besides the statistical figures, there are qualitative studies on the migration of Haitian people to Brazil. Here I will briefly describe 4 papers on the subject. First, de Moraes, de Andrade, and Mattos (2013) present a historical, social and political summary of the crisis in Haiti, since its Independence until 2013, and of the legal and social aspects and challenges involving the intense immigration from Haiti. They highlight the difficulties in getting a humanitarian visa to Brazil, in avoiding insecure and dangerous jobs, and in securing access to civil and human rights. Although Brazil has been offering support and welcoming migrants from Haiti, there still is work to do in relation to the training of security officers at the country's borders and in relation to social assistance focused on this population. Gomes (2017) discusses the issue of social assistance in more detail.

In this second paper, Magalhães and Baeninger (2017) interviewed 31 Haitian people living in Balneário Camboriú, Santa Catarina, to describe this sample of the population and their migration routes. Magalhães and Baeninger asked questions about the trajectories they took when leaving Haiti and entering Brazil, about their main goals at the moment (e.g. bringing relatives to Brazil, sending money to their family who stayed in Haiti), and about their interests for the future (e.g. going back to Haiti, staying in Brazil, moving to a different country). They also asked their ages and occupations. Magalhães and Baeninger conclude that Haiti has been historically marked by emigration and that most Haitian immigrants in Brazil are interested in providing stability to their families. Migrating is, then, not simply a personal choice, it is a strategy for the family due to an economic and social crisis.

The third paper offers a dense social description. Gomes (2017) was interested in the migration trajectories and the ways of living of Haitian immigrants in Florianópolis and surrounding cities. She highlights that, although Brazil has a national law (Law number 9,474/1997, National Law for Refugees, Lei Nacional para Refugiados) and a national council (National Council for Refugees, Conselho Nacional de Refugiados, CONARE) specific for the support of refugees, there are still many reports of violation of human rights (Gomes, 2017). Gomes presents qualitative results from interviews with Haitian people living in Florianópolis and nearby cities. She carried out these interviews for one year; however, she does not mention the total number of participants. Gomes organized the reports in two categories: reports from university students and from workers. The most pressing reasons for migrating seem to be different between these groups. Students talked more about their desires and ambitions, while workers mentioned the obligation to help their families have a better life. She points out that, for this sample of the population, economic factors are the ones that motivate emigration from Haiti the most, but that they are not the only ones. For example, participants commented that access to the university would be one of the best cenarios. In addition, the majority of interviewees thought Brazilians were welcoming and friendly; however, none of the participants had a closer relationship with a Brazilian. This loneliness was also observed by Rodrigues (2021).

Rodrigues (2021) interviewed a first-year Haitian student at a public university in Brazil. She explains that he arrived in Brazil in 2017 to study Geography. This major was not his first choice; it would have been Social Sciences, the undergraduate course he was enrolled into back in Haiti. However, there was no vacancy, and he had to choose a different course. He reported situations involving racism and other prejudices during his first year at the university. In addition, he commented that being highly proficient in Brazilian Portuguese was very important, that he was doing his best to develop his linguistic skills, and that he felt that was able to express and do much more in his native language than in BP and he would like to show what he was capable of. Rodrigues concludes that, even with civil rights for refugees, this Haitian student still suffers silencing, prejudices, and exclusion.

Historically, Haiti is scarred with emigration waves. A high flux of emigrants from Haiti is considered common, and their preferred destiny is the United States (Magalhães; Baeninger, 2017). Now Brazil could be added as a next preferred country, probably because of the following attractive aspects: economic growth, sports, culture, and welcoming people — all mentioned by the interviewees of de Moraes, de Andrade, and Mattos (2013).

At this point, before presenting some of the structure of the HC language, it is important to define a few linguistic and sociopolitical concepts related to bilingualism and multilingualism here, which are majority, minority, heritage, and welcoming languages.

Montrul (2015) defines a majority language as usually being recognized as the official language of a country, used by means of communication, and employed in educational, administrative and governmental spheres, such as BP in Brazil. Meanwhile, a minority language is used by ethnolinguistic minority groups and generally is marginalized due to social, political and cultural factors. This is the case of indigenous and heritage languages in Brazil. Altenhofen (2013, p. 94, my translation) defines minority languages as "the modality of languages or varieties used at the margins or in parallel to the dominant (majority) ones", and states that the key characteristic of a minority language is its political status.

Then, a heritage language may be defined as a language, other the official language(s) and aboriginal ones, spoken by immigrants and their children or may be defined as "minority languages co-existing with majority languages, including immigrant languages, national minority languages, and aboriginal languages" (Montrul, 2015, p. 15). In Brazil, a heritage language entails not only linguistic skills, but also the family and the original country's identity and culture (Brasil, 2023). Flores and Melo-Pfeifer (2014) define heritage language as the one spoken by the migrating family. Meanwhile, Barbosa, Fistarol, and Silveira (2020) provide separate definitions for immigration and heritage languages. For them, an immigration language is the one used by an immigrant in a different country while a heritage language is the one used by the people born in this country in immigrant families.

Montrul (2015) also conceptualizes heritage speaker. She explains that the heritage speaker is a bilingual who grew up exposed to a minority language usually at home and to a majority language other circles in the society. It is emphasized that this exposure leads to proficiency in the minority language. Thus, the heritage speaker presents some degree of competence in the heritage language. However, this definition may vary according to local sociolinguistic aspects.

Lastly, a welcoming language may be viewed as an extension of an additional language in that it is added to other languages of the speaker with no hierarchy (Ferreira, 2021). Teaching a welcoming language implies that the reasons and contexts for learning this language should guide every step of the process, so intercultural characteristics are taken into consideration in order to promote the understanding of sociocultural practices and the migrants' autonomy to act in this new society with dignity (Modesto-Sarra, 2022; Rocha; da Silveira Gileno, 2022). Thus, the social, cultural and political impacts of teaching and learning a welcoming language are clearly stated.

Despite the relevance of the debate behind these definitions, and especially due to the implications of choosing one concept or the other, I will refrain from adopting them in this dissertation. I will refer to BP as the participants' second language (L2) since it is not their native language and is key to the functioning of institutional and social aspects of a community or country (Miao, 2015). In addition, participants reported having diverse language history backgrounds. For example, 26 of them mentioned having learned French and/or Spanish before moving to Brazil. Thus, BP could be classified as L2, L3 or even L4 depending on the participant. Finally, according to the definitions above, participants could be considered speakers or minority, heritage or immigration languages. In relation to the Brazilian society, BP is indeed the majority language. However, in this dissertation, BP will be referred to as the L2 in the case of HC-speakers living in Brazil.

Now I will present a very brief description of the structure of the HC language. Although I will focus on processing at the lexical level in both studies described here, I add this section about the structure of HC so that the influences from other languages may be visualized.

2.3.2 Language structure

Albert Valdman offers a thorough description of HC and of the research on this language conducted since 1978 in his book "Haitian Creole: Structure, Variation, Status, Origin", published in 2015. I could not find any digital version of the book, and the physical version costs around R$300.00. However, I could access a review of the book written by Sibylle Kriegel (2016). Valdman explains that the phonological system of HC is very similar to the French one and that since 1940 orthographies had been developed for HC, one of which was declared as the official one (Kriegel, 2016, p. 330). It is also stated that "Haitian Creole is capable of meeting all the communicative needs of a language on its way to literacy" (Kriegel, 2016, p. 330). In addition, Kriegel summarizes some typical basic sentence characteristics of HC presented by Valdman, such as

> the relative lack of inflectional affixes, the multifunctionality of word classes, and the resulting importance of word order. While these features are shared by almost all creoles, the following section on the copula deals with a feature more specific to Haitian Creole, the syntactic operators *ye* and *se*. As in numerous other French-based creoles, reflexivity can be marked either by concepts referring to the human body plus the personal pronoun (*kò*/*tèt* + pronoun) or by a direct-object pronoun, and there is no overt passive marking (Kriegel, 2016, p. 330).

In a broad way, the origins of Haitian Creole are especially Fongbe, a Niger-Congo language, and French. Most phonological representations and lexical items in HC seem to come from French, the superstratum language, while many syntactic and semantic properties seem to derive from Fongbe, the substratum language, and related languages. Lefebvre (1998) describes the differences and similarities across HC, Fongbe and French and proposes that the former one was influenced by the last two languages. According to Lefebvre (1998) research, here are some characteristics of HC.

In HC, a nominal phrase (NP) can be formed by a noun followed by a possessor phrase, a demonstrative term, a definite marker, and/or a plural marker. All nominal constituents are postnominal. A relative clause may be inserted between the noun and the determiner.

(1)	krab	[mwen ø]		sa	a	yo
	crab	me	GEN	DEM	DET	PL

‘these/those crabs of mine (in question/that we know of)’
(Lefebvre, 1998, p. 78).

The determiner is not marked for gender and may take the forms *la*, *a*, *an*, *nan*, and *lan*. It also may be covert, that is, HC allows for bare NPs. In addition, the determiner is essential for clause structure and has a role in relative, conditional and factive clauses.

The plural marker *yo* is marked as [+definite] and is also the 3rd person plural pronoun. On the other hand, singular has no specific marker.

(2)	Yo	pati.
	3rd-PL	leave

‘They left.’
(Lefebvre, 1998, p. 85).

HC has two demonstrative forms, *sila*, which is marked as [−proximate], and *sa*, which may be either [−proximate] or [+proximate]. These deictic forms occur within NPs and may replace a NP which is a verb argument.

(3)	M'	wè	–	sa	//	sila.
	I	see	one	[+deic]	//	[−prox]

'I saw this/that // that one.'
(Lefebvre, 1998, p. 90)

Determiners in HC can co-occur in the NP and can have a phonologically null head. Moreover, they are not marked for gender, number or animacy.

HC has no overt case markers. There is a Genitive phrase, composed of a demonstrative term, a determiner and a plural marker. However, there is no overt Genitive marker.

(4)	timounn	Mari	a
	child	Mary	the

'(the) Mary's child'
(Lefebvre, 1998, p. 106)

HC has a tense, mood and aspect system in which tensed verbs always occur in their simple form. There are no markers for person or number and no affixes. However, there are markers between subject and verb for encoding time, mood and aspect.

(5) Anterior:
· Past/past perfect: *te*

Irrealis:
· Definite future: *ap*
· Indefinite future: *a-va, va, av, a*
· Subjunctive: *pou*

Non-complete:
· Habitual: –
· Imperfective: *ap*

(Lefebvre, 1998, p. 112, only HC markers)

In HC, all tenses are relative to a reference point. The definite future expresses that the speaker is certain that an event will happen while the indefinite future conveys the idea that an event has some chance of happening. The subjunctive could be translated into English as *must*, *should* or *may*. Some speakers allow that the definite future *ap* and the imperfective *ap* occur in sequence in a sentence. Others do not permit this repetition but still keep the semantic interpretation. The markers above can be combined to form complex verb tenses.

(6) M' ap ap sòti.
I DEF-FUT IMP go-out
'I will be going out.'
(Lefebvre, 1998, p. 112)

(6) M'	ap	ap	sòti.
I	DEF-FUT	IMP	go-out

In HC, verbs may have a dynamic (event in progress), resultative (result of an event) or stative (not a process) aspect. The temporal interpretation of a sentence depends on which marker co-occurs with which aspect. For example, the tense marker *te* followed by a stative verb will always be interpreted as past perfect. On the other hand, the marker *te* followed either by a dynamic or a resultative verb may express past or past perfect. The imperfective marker *ap* followed by a dynamic verb will always express an ongoing event while *ap* followed by a resultative or stative verb cannot be interpreted as so. The definite future marker *ap* and the indefinite future marker *a-va* occur with all verb aspects; they may express future perfect, and conditional when coupled with *te*. Finally, the mood marker *pou* can express a wish, an obligation, or an order.

Moreover, bare sentences, that is, sentences with no marker before the verb can have their tense expressed through specific elements. They are the aspect of the verb and the definiteness of the subject and/or the direct object. For example, present perfect is manifested as a dynamic verb followed by a definite object. On the other hand, a dynamic verb followed by an indefinite object can mean past or general present. Resultative verbs may indicate present or present perfect and stative verbs always express present.

In HC, personal pronouns are marked for person and number but not for gender. Strong forms are the following:

(7) mwen	1st-person singular	'I, me'
ou / [wu]	2nd-person singular	'you (sg.)'
li	3rd-person singular	'he/she/it /him/her'
nou	1st- and 2nd-person plural	'we/us/you (pl.)'
yo	3rd-person plural	'they/them'

(Valdman *et al*. 1981 apud Lefebvre, 1998, p. 141)

Weak forms are as follows:

(8)	Strong forms	Weak forms
1st sg.	mwen	m
2nd sg.	ou	u / w
3rd sg.	li	l
1st, 2nd pl.	nou	n
3rd pl.	yo	y

(Lefebvre, 1998, p. 151)

Strong forms occur before and after consonants while weak forms precede vowels.

Only strong forms can be conjoined, cleft and topicalized. Weak forms may be in NPs and may complement a preposition. There are no possessive adjectives; and possession is expressed through Genitive phrases. For example,

(9) Jan pran [pa [Mari / mwen ø]].
John take thing Mary me GEN
'John took Mary's/mine.'
(Lefebvre, 1998, p. 146)

Reflexivity is expressed via personal pronouns in the object position and specific nouns. For example, "*tèt* head and *kò* body followed by a possessor phrase containing a personal pronoun" (Lefebvre, 1998, p. 159) can be interpreted as reflexive.

(10) M_i wè m_i.
I see me
'I saw myself.'
(Lefebvre, 1998, p. 160)

Alternatives would be the items pronoun+MÈM, which can also be used for emphasis, or *tèt-li*.

(11) Jan_i pale ak $li_{i/j}$ / tèt $-li_i$.
'John speaks with him/ himself.'
(Lefebvre, 1998, p. 165)

The following are the Wh-words in HC.

(12) ki-lès 'which one'
(ki-)sa 'what'
kouman 'how'
konbyen 'how much, how many'
(Brousseau, 1995 *apud* Lefebvre, 1998, p. 172)

And these are the Wh-phrases:

(13) ki mounn 'which person/who'
ki bagay 'which thing/what'
(ki) kote / ki bò 'which place/where'
ki jan 'which manner/how'
ki kalite 'which kind/how'
ki lè 'which time, moment/when'
pou ki(-sa) 'for what/why'
(Lefebvre, 1998, p. 173)

In addition, HC presents phonologically null complementizers in sentences with verbs "such as *kwè* 'believe', *di* 'say', *panse* 'think'" (Lefebvre, 1998, p. 184). As seen in the Wh-phrases above, *ki* substitutes the subject.

(14) Jan kwè / di / panse [CP ø Mari vini].
John believe / say / think COMP Mary come
'John believed/said/thought that Mary came.'
(Lefebvre, 1998, p. 184)

The complementiser *Pou* follows some verbs and adjectives such as *want* and *good* respectively. It has the same pronunciation as the preposition *pou* and the mood marker *pou*.

(15) Yo te vle [pou m te antre nan troup Jakmèl].
they ANT want COMP me ANT join in troops Jacmel
'They wanted me to join Jacmel's troops.'
[Lit.: 'They wanted that I joined Jacmel's troops.']
(Koopman; Lefebvre, 1982 *apud* Lefebvre, 1998, p. 188)

Relative and factive clauses are introduced by an operator. However, it is phonologically null in HC.

(16) Rive ø Jan rive a...
arrive OP John arrive DET
'The fact that John arrived...'
(Lefebvre, 1998, p. 204)

Clauses are joined using the conjunction *(e)pi*.

(17) Jan pati (e)pi Mari rive.
John left and Mary arrived.'
(Lefebvre, 1998, p. 205)

NPs can be conjoined by the preposition *ak* or *(kòl)ak* (from *kòle-ak*, which means *close with*).

Se is used in HC cleft sentences in the beginning of the constituent. There must be a nominal predicate for *se* to appear.

(18) Se Jan Mari wè.
(...)
'It is John that Mary caught sight of.'
(Lefebvre, 1998, p. 206)

In HC, the negation marker pa usually precedes all verb markers. It must appear after the subject and before the verb.

(19) Jan pa t' av- ale nan mache.
John NEG ANT IND-FUT go in market
'John would not have gone to the market.'
(DeGraff 1993 *apud* Lefebvre, 1998, p. 208)

Moreover, there is no marker of yes-no questions at the end of sentences. Rising intonation or *èske* at the beginning are used.

(20) Èske Jan vini?
Q John come
'Has John come?'
(Lefebvre, 1998, p. 212)

HC has a marker of insistence or divergent opinions across speakers. It is *non*; when it appears in affirmative sentences, they must also be imperative.

(21) Pa ale non!
NEG GO INS
'Don't go!'
(Lefebvre, 1998, p. 215)

According to this brief description of some characteristics of HC, it is possible to observe its relation to other languages. Lefebvre (1998) mentions that many phonological representations of items come from the superstratum language, in this case, French, while most semantic and syntactic properties seem to have originated from the substratum language, that is, Fongbe.

In this section, I presented a brief overview of the structure of HC. In the next section, I will list and summarize studies which investigated HC.

2.3.3 Studies about HC

In January 2020, I carried out a search on CAPES Journal Library using the keyword "Haitian Creole" in quotes. The search returned 1,773 results for peer-reviewed studies, which is a small set of studies. Among them, 654 were organized in the following categories: Linguistics, Psychology, Bilingualism, Language and Literature, and Second Language Learning. From these works, I selected 5 studies to describe in my qualification defense in March 2020. Only two of them (Archer *et al.*, 2018; Cazeneuve; Nascimento, 2016) appeared again when I rerun this search three years later. I will describe the 5 of them here because they are still relevant for this work.

Using an Education account, DeGraff (2016) discusses the importance of the use of a native language in the formal education of children, especially in their first years of schooling. He advocates for the actual implementation of HC, one of the official languages of Haiti, in schools for improving levels of literacy and education in the country. Haiti was a French colony until 1804. All of its population speaks HC and only 5% speak French, both official languages. However, all Haitian legislation, official documentation, press news, and formal instruction are carried out in French — despite the prescription that HC should be used. Consequently, pedagogical materials in HC are lacking, most of the population is illiterate. He argues that, according to UNESCO, education should be conducted in mother tongues so that it may be of high quality and of optimal access.

DeGraff was part of the project "Mother Tongue Books: Learning to Read in Haiti", led by Christine W. Low, which had three phases: a baseline assessment (pretest, 2012-2013), an intervention (treatment, 225 students, 2013-2014), and an evaluation phase (posttest, 2014). They used 8 of the 9 subtests of the Early Grade Reading Assessment (EGRA) instrument in the pretest and in the posttest. Their intervention was to create pedagogical materials in HC with child-centered interactive pedagogy in five Haitian schools; their control group was a school (Lekòl Kominotè Matènwa, LKM) which already used this approach. Overall, in the pretest, students from the LKM school scored significantly better at EGRA than students from non-LKM schools, who had really low scores. In the posttest, it is possible to observe a positive effect from the treatment phase, and students from the non LKM schools substantially improved their performance in relation with their peers from the LKM school. Thus, the author emphasizes the crucial role of the mother tongue in developing literacy.

In line with debates on language teaching, Cazeneuve and Nascimento (2016) described the bilingual status of Haiti and its impact on education. The country was declared independent in 1804, and the Independence Act was written in French, ignoring the fact that the HC was also spoken. Only 2% of the population could speak and write, especially in French. Thus, the elite controlled the country and the law through a linguistic symbol of power. In 2012, more than 80% of the Haitian population worked with agriculture, either professionally or not. It is only through learning French, which is the foundation for formal education, that a person may improve their life. Basic formal instruction lasts 9 years, and at the end of the last year of high school, students need to pass an exam in order to receive their diploma. In 2011, 88% of schools were private ones. Pedagogical resources for developing literacy and also for learning French are lacking. In 1979, HC was indicated as the language of teaching during the first years of schooling. The idea was to change the structure of the Haitian educational system completely.

However, teacher education and pedagogical materials were obstacles. Consequently, around 500 thousand children are not able to write using Creole and are not able to speak or write using French.

Under the lenses of Speech Production research, Archer *et al* (2018) studied the typical language development of HC-speaking children, focusing on phonological acquisition. Their objective was to identify the phonemic inventories used by 2- to 4-year-old Haitian children so that speech language pathologists may have a baseline for diagnosis. Twelve children — 8 from low socioeconomic status, 3 from middle, and 1 from high — participated in the study. The examiner used toys, colored pictures and picture books to engage the children in activities and name (or repeat) 89 stimuli. Consonants produced by the participants were categorized and frequent substitutions were annotated. Two-year-olds seemed to have acquired 6 phonemes; 3-year-olds, 13; and 4-year-olds, 19 out of 23. Posterior phonemes seemed to be the most difficult ones to accurately produce. These results are in accordance with studies on HC and Canadian French acquisition milestones.

Using an account from Communication Studies, Léger and Armbrister (2009) investigated the factors affecting students' perceptions of HC in the city of New Provence, The Bahamas. They explain that, although people still think of HC as a depreciated variety of French, it is indeed a language distinct from French and "the most widely spoken Creole language in the world" (Léger; Armbrister, 2009, p. 3). The lexicon of Creole derives from French and many other languages; however, morphology, syntax and pronunciation are very different. For example, possessive pronouns are placed after words, verb tenses are marked using particles, and it is possible to place emphasis syntactically. The authors mention that Haitians tend to immigrate, mostly illegally, to The Bahamas, seeking a better life. Thus, the social and political situation in Haiti and in The Bahamas promotes prejudice against HC and its speakers. Despite that, The College of The Bahamas (COB) offers HC as one of the languages students have to choose when studying foreign languages. Since the reasons for a student to select HC are not the same as for Spanish or French, foreign language teaching techniques are a little different. The authors carried out a survey with 419 randomly selected students from COB and other schools, which showed that generally students in New Provence agree that HC should be taught and studied, but that it is "broken French" and it should not be the second language of The Bahamas. Students also mostly perceive The Bahamian dialect negatively. The authors conclude that there may be other factors influencing students who think negatively of HC into choosing to study this language.

Finally, from a more Sociolinguistic point of view, Pereira and Costa (2015) looked into language contact between Brazilians and Haitian immigrants in Rio de Janeiro. They describe the diglossia existent in Haiti which grants higher prestige to French over HC despite the latter being spoken by the majority of the population and the former only by an elite. French is also used in legal and administrative spheres, and in all formal education, while HC is used in informal and family-related situations. Haitians emigrated mostly due to socioeconomic and political circumstances of their country. Since 2010, Brazil has received immigrants from Haiti, and 30% of them are now working in civil construction. The authors present a historical overview of immigration laws since 1850 until today. Then, they mention that it is possible to understand social phenomena via the study of contact between languages and the engaged social networks. Dense networks tend to discourage linguistic change due to little access to novelties and strong social norms. The authors asked that 16 Haitians students of civil construction completed a questionnaire on their social profile and language use. Results show that this sample of immigrants still uses HC and French to communicate with family in Haiti and HC to communicate with friends in Brazil. However, their social networks are changing and include Brazilian Portuguese, which is important in developing their profession and education. French is an international bridge and refers to formal and academic instances.

Three years later, in October 2023, I repeated the search, and the number of studies available on CAPES Journal Library was even smaller. There were 429 results for "Haitian Creole". Among these, 246 were labeled as peer-reviewed, and 214 were articles. Part of these 214 was organized as follows: 39 from Linguistics, 26 from Language, and 24 from Language & Linguistics. However, only one of the 5 studies I described in my qualification was available in this search. I will add the other 4 studies as theoretical background. Among the 39 studies from Linguistics, 6 of them only mentioned HC and were not interested specifically in it, 1 was from Literary studies, and 1 appeared twice on the list. Among the 26 studies from Language, 10 also appeared under Linguistics, 1 was from Literary studies, 1 was a grammar book, 2 were not specifically about HC, and 11 were mainly associated with areas such as Medicine, so only 1 result was new. Among the 24 studies from Language & Linguistics, only 1 was new, that is, had not appeared in the other 2 categories. I rerun the search in CAPES Journal Library using the BP equivalent of the keyword, "crioulo haitiano" in quotes, and I got 14 results. Only 6 were peer-reviewed, and 2 of them were from Literary studies. 37

Then I replicated the search on the Web of Science database. "Haitian Creole" returned 187 results, of which 130 were articles. Among these, 52 belonged to Language & Linguistics and 45 to Linguistics. Among the 52 from Language & Linguistics, 13 had already appeared in previous searches, and 5 were not specifically about HC. Among the 45 from Linguistics, only 2 studies were new results in the searches. The keyword "crioulo haitiano" returned no results. All the results from the searches from October 2023 are presented in Table 1.

Among the 74 studies I could find, Hebblethwaite (2010) presents the one with approaches more similar to the ones used in Psycholinguistic investigations. Benjamin Hebblethwaite (2010) investigated how adverbs are code-switched across HC and English and discussed the results in relation to the definition of adverbs in the discourse and the classification of some occurrences as actual code-switches or as either borrowings-in-progress or borrowings. His theoretical background was anchored in the Minimalist Program and his method had similar steps as some used in Corpus Linguistics and in Psycholinguistics. He argued that code-switching is always bidirectional and that "the lexical and functional properties of words and the sociolinguistic and psycholinguistic realities of the community provide a better framework for predicting outcomes in code-switching" (Hebblethwaite, 2010, p. 410). In addition, he expected that code-switching patterns would show that adverbs could be organized into two types of members: lexical ones (e.g. fast) and functional ones (e.g. very), which can co-occur (e.g. very fast). In this study, Hebblethwaite used two corpora from 2004 and 2005 about naturalistic interviews with second-generation Haitian Creole-English bilinguals. He coded all tokens in both transcribed corpora according to the language of the token. This way, he could observe the direction of the instances of code-switching. For example, an instance of a word in HC, then a switch to English, and a return to HC would be coded as xYx, in which x represents a token in HC, y a token in English, and capital letters represent the target token. His results show that most (66.7%) tokens of adverbs in the corpus from 2004 were from HC unmixed contexts, that is, where there was no code-switching during speech in HC; 19.8% were tokens from unmixed English contexts, that is, where there was no code-switching during speech in English. This means that 86.5% of adverb occurrences in this corpus happened in no-switch contexts. Despite that, Hebblethwaite considered the 12% of English adverbs during HC speech as a "striking percentage" (Hebblethwaite, 2010, p. 418) in comparison with the 1.27% of HC adverbs during English speech. He pointed out that the code-switching instances that are repeated the most in the corpus (e.g. "(e)specially, maybe, still and then" (p. 419)) are probably borrowings-in-progress, not code-switches anymore. However, he explained that the hypothesis that adverbs could be organized into two types of members — lexical ones and functional ones

Table 1 – Works about HC vocabulary and language structure

Authors	Year	Title	Linguistics area
Tinelli, Henri	1974	Generative and creolization processes: Nasality in Haitian Creole	Phonetics
Bentolila, A.	1987	Aspecto-temporal marks in Haitian Creole, from a synchronic analysis to the formulation of diachronic hypotheses	Historical
Cadely, Jean-Robert	1988	Phonological studies in Haitian Creole - A resolved paradox	Phonetics
Ndayiragije, J.	1989	The source of the agglutinated determinant in Haitian Creole	Syntax
Lefebvre, Claire; Brousseau, A. M.; Filipovich, S.	1989	Haitian Creole morphology - French phonetic matrices in a West-African mold	Morphology
Damoiseau, Robert	1989	Communication and functioning of Haitian Creole - An analysis	Syntax
Lumsden, John S.	1990	The biclausal structure of Haitian clefts	Syntax
Singler, J. V.	1990	On the use of sociohistorical criteria in the comparison of creoles	Historical
Byrne, F.	1990	Toward an account of preclausal focus in some creole languages	Syntax
Lumsden, John S.; Lefebvre, Claire	1990	Predicate-cleft constructions and why they aren't what you might think	Syntax
Deprez, Viviane	1992	Raising Constructions in Haitian Creole	Syntax
Deprez, Viviane	1994	Haitian Creole - A pro-drop language	Syntax
Lumsden, J. S.	1994	Possession - Substratum semantics in Haitian Creole	Semantics
Damoiseau, Robert	1994	Reflection on the functioning of the system of verb aspects and tenses in Haitian Creole	Syntax
Lumsden, J. S.	1995	Aspect and lexical semantic representations in Haitian Creole	Semantics
Lefebvre, Claire	1996	The tense, mood and aspect system of Haitian Creole and the problem of transmission of grammar in Creole genesis	Syntax
Singler, J. V.	1996	Theories of Creole genesis, sociohistorical considerations and the evaluation of evidence: The case of Haitian Creole and the relexification hypothesis	Historical
Deprez, Viviane; Vinet, M. T.	1997	Predicative constructions and functional categories in Haitian creole + Linguistics, aspect, sentence structure	Syntax

Lefebvre, Claire	1997	Relexification in creole genesis + Language formation: The case of demonstrative terms in Haitian creole	Syntax
Zephir, F.	1997	The Social Value of French for Bilingual Haitian Immigrants	Social
DeGraff, Michel	1999	Empirical quicksand: Probing two recent articles on Haitian creole	Critics
Lefebvre, Claire	1999	Substratum semantics in the verbal lexicon of Haitian creole (French, Fongbe)	Semantics
Lumsden, J. S.	1999	The role of relexification in creole genesis	Syntax
Nikiema, E	2000	Lexical and epenthetic initial vowels in Haitian creole	Phonetics
Lefebvre, Claire	2000	On data (Linguistic analysis)	Critics
Frederking, Robert; Rudnicky, Alexander; Hogan, Christopher; Lenzo, Kevin	2000	Interactive Speech Translation in the Diplomat Project	Computational
Lefebvre, Claire	2001	On the semantic opacity of creole languages	Semantics
Lefebvre, Claire	2001	The interplay of relexification and levelling in creole genesis and development	Morphosyntax
Cadely, Jean-Robert	2002	Status of nasal vowels in Haitian Creole	Phonetics
DeGraff, Michel	2003	Against Creole Exceptionalism	Historical
Valdman, A	2005	Towards a bilingual Haitian Creole-French/French-Haitian Creole dictionary aimed at school pupils	Educational
Fattier, D	2005	Comments on the questioning of Haitian creole	Critics
Mather, Patrick-Andre	2006	Second language acquisition and creolization - Same (i-) processes, different (e-) results	Syntax
Fitzpatrick, Justin M.	2006	Deletion through Movement	Syntax
Bonet, Eulàlia; Lloret, Maria-Rosa; Mascaró, Joan	2007	Allomorph selection and lexical preferences: Two case studies	Morphology
Takahashi, Shoichi; Gračanin-Yuksek, Martina	2008	Morphosyntax of Movement Dependencies in Haitian Creole	Morphosyntax
Harbour, Daniel	2008	Klivaj predika, or predicate clefts in Haitian	Syntax
Guijarro-Fuente, Pedro; Lopez, Luis A. Ortiz	2008	Creole/Spanish contact and the acquisition of clitics on the Dominican-Haitian border	Morphosyntax

Hebblethwaite, Benjamin	2009	Scrabble as a tool for Haitian Creole literacy Sociolinguistic and orthographic foundations	Educational
Hebblethwaite, Benjamin	2010	Adverb code-switching among Miami's Haitian Creole-English second generation	Psycholinguistics
Yavas, Mehmet	2011	Patterns of cluster reduction in the acquisition of #sC onsets: Are bilinguals different from monolinguals?	Phonetics
Farr, Marcia	2011	Urban plurilingualism: Language practices, policies, and ideologies in Chicago	Political
Kleyn, Tatyana; Reyes, Sharon Adelman	2011	Nobody said it would be easy: ethnolinguistic group challenges to bilingual and multicultural education in New York City	Educational
Hebblethwaite, Benjamin	2012	French and underdevelopment, Haitian Creole and development Educational language policy problems and solutions in Haiti	Political
Lefebvre, Claire	2013	A comparison of the nominal structures of Saramaccan, Fongbe and English with reference to Haitian Creole: Implications for a relabelling-based account of creole genesis	Syntax
Costa-jussà, Marta R.; Banchs, Rafael E.	2013	Automatic normalization of short texts by combining statistical and rule-based techniques	Computational
Schwarz, Florian	2013	Two Kinds of Definites Cross-linguistically	Semantics
Glaude, Herby; Zribi-Hertz, Anne	2014	Verb cognates in Haitian Creole	Syntax
McWhorter, J.	2014	Saramaccan and Haitian as young grammars: The pitfalls of syntactocentrism in creole genesis research	Historical
Valdman, Albert; Villeneuve, Anne-Jose; Siegel, Jason F.	2015	On the influence of the standard norm of Haitian Creole on the Cap Haitien dialect Evidence from sociolinguistic variation in the third person singular pronoun	Historical
Hebblethwaite, Benjamin	2015	Corrections on the History and Design of Haitian Creole Scrabble	Educational
Carter, Phillip M.; Lynch, Andrew	2015	Multilingual Miami: Current Trends in Sociolinguistic Research	Social

Obata, Miki; Epstein, Samuel; Baptista, Marlyse	2015	Can cross-linguistically variant grammars be formally identical? Third factor underspecification and the possible elimination of parameters of UG	Syntax
Cazeneuve, Miseline; Nascimento, Lilian Cristine Ribeiro	2016	A influência da situação linguística do Haiti no processo de aprendizagem das crianças	Political
DeGraff, Michel; Stump, Glenda S.	2018	TEACHING LINGUISTICS & LANGUAGE AND PUBLIC POLICY Kreyol, pedagogy, and technology for opening up quality education in Haiti: Changes in teachers' metalinguistic attitudes as first steps in a paradigm shift	Educational
Archer, Justine; Champion, Tempii; Tyrone, Martha E.; Walters, Sylvia	2018	Phonological Development of Monolingual Haitian Creole–Speaking Preschool Children	Phonetics
Machry da Silva, Susiele; Bassols Brisolara, Luciene	2018	Teaching of Portuguese to speakers of other languages: an analysis of transfers from the mother tongue into writing	Educational
Machry da Silva, Susiele; Bassols Brisolara, Luciene	2018	Ensino do Português para Falantes de Outras Línguas: Análise das Transferências dos Padrões da LM na Escrita	Educational
Deprez, Viviane	2019	Plurality and definiteness in Mauritian and Haitian creoles	Semantics
Govain, Renauld	2019	L'état des lieux du créole dans les établissements scolaires en Haïti	Educational
O'Neil, David	2019	The Middle English Creolization Hypothesis: Persistence, Implications, and Language Ideology	Historical
Foltran, Maria José; Rodrigues, Patrícia; Bertucci, Roberlei Alves; Deschamps, Thais	2019	The 12th Workshop on Formal Linguistics: Foreword	Social
Damoiseau, Robert	2019	Pour une approche comparative de la grammaire créole : Créoles guadeloupéen, martiniquais, guyanais, haïtien	Educational
Teixeira, Lovania Roehrig; Machry da Silva, Susiele	2019	Morfologia flexional dos verbos no Português Brasileiro e no Crioulo Haitiano	Morphology
Silva, Adelaide Hercília Pescatori	2019	Apontamentos para o ensino do sistema sonoro do PB	Educational

Sumonte, Valeria	2020	Desarrollo de la competencia comunicativa intercultural en un programa de adquisición de la lengua criollo haitiana en Chile	Educational
Seguin, Luisa	2020	Transparency and language contact The case of Haitian Creole, French, and Fongbe	Historical
Ulysse, Gerdine M.; Al Masaeed, Khaled	2021	The influence of socio-economic status, age, gender, and level of literacy on language attitudes The case of Haitian Gonaviens	Social
Lacoste, Véronique	2021	Mediating linguistic diversity in the diaspora: An illustration from Haitian Canadians	Social
Joseph, Lefranc	2021	Lang nan pwodiksyon sosyete ayisyen an: Istorik, fòs pouvwa akkonsyantizasyon	Social
Bartens, Angela	2021	The making of languages and new literacies: San Andres-providence creole with a view on Jamaican and Haitian	Historical
Tezil, David	2022	On the influence of Kreyol swa Evidence from the nasalization of the Haitian Creole determiner/la/in non-nasal environments	Phonetics
Nunes, Ariele Helena Holz; Agostinho, Ana Livia	2022	Structural properties of ideophones in Kreyol	Morphophoneti cs

Source: The Author (2024)

— was not confirmed based on the corpus. So, he concluded that “adverbs behave like a lexical category” (Hebblethwaite, 2010, p. 419). He extended the discussion to left-periphery elements, borrowings, and verb/adverb interactions.

This search in the literature indicated that there was no study about HC I could find which was conducted according to methods used in Psycholinguistics. This was one of the gaps in the literature I intended to contribute to filling. In the next section, I will present the existent literature on the relationship between reading habits and language development.

2.4 READING HABITS

When decoding skills are automatized, the reader may read with the objective of learning. This includes learning new words, especially when the constituents of known words are of high quality, that is, fully specified (Perfetti, 2007). Thus, in theory reading habits would have an impact on vocabulary size and reading skills. Would it be implausible to extend this assumption to L2 reading constructs? The literature seems to indicate that it may be conceivable that L1 and L2 reading habits are associated with L1 and L2 language processing. First, I will present studies which test this possibility only within languages, then I will present the ones which test it between languages.

Considering frequency of reading in one language, Rudell and Hu (2010) investigated how the recognition of words and of Gestalt figures was influenced by reading experience. Bilinguals, with low experience in the L2, and monolinguals completed a task in which they had to recognize English words and Gestalt figures. Monolinguals were faster to answer to words and figures than bilinguals and showed shorter ERP latencies for words than bilinguals. Bilinguals' responses to L2 words were longer than to Gestalt figures; however, there was no difference in ERP latencies for figures between groups. Monolinguals were much more exposed to English than bilinguals and therefore had higher experience with the language. Thus, generally speaking, higher experience with words contributes to higher speed of lexical access.

Yuko Butler (2011) investigated the influence of background factors in the *kanji* reading and writing skills of 4th graders. She invited 13 native speakers of Japanese and 27 second language learners. The L1s of the non-native speakers were minority languages while Japanese was the majority one. Participants completed *kanji* reading and writing tests and filled a linguistic background questionnaire. Butler observed that, although the L2 learners had equivalent oral proficiency levels in comparison with the native speakers, they showed more difficulties in *kanji* reading. This difference in reading skill was not influenced by simply being a native speaker of Japanese, since oral proficiency was similar across groups. Kanji reading was actually significantly and positively impacted by "the frequency of reading Japanese books outside of school" (Butler, 2011, p. 14), that is, reading more Japanese books outside of school was associated with better *kanji* reading skills. However, there was no correlation between reading habits in Japanese and in other languages for this sample of children. This indicates that the frequency of reading in one language does not necessarily influence frequency of reading in the other language for this group.

With regards to the fact that language development is affected by linguistic habits, Pratheeba and Krashen (2013) tested the relationship between L2 reading experience, operationalized as reading habits, and L2 vocabulary. Also, they verified the influence of online reading materials on vocabulary. They prepared a 20-question questionnaire about reading habits in English and a vocabulary test for Indian students, who spoke three L1s and English as an L2, to complete. Answers for the questionnaire were given according to a 6-point scale, where 0 stood for "never" and 5 for "always". Performance in the vocabulary test and answers to the questionnaire correlated positively and significantly. The book reading question subset presented a moderate positive correlation with vocabulary. However, only one item from the computer reading subset (reading for pleasure) was positively associated with vocabulary. Thus, reading experience is important to language development since better scores on the vocabulary test were related to higher frequency of reading. A vaster vocabulary allows more frequent reading habits, which in turn, lead to an increase in vocabulary (Pratheeba; Krashen, 2013).

Sun, Bornstein and Esposito (2021) tested 736 children in year 1 of kindergarten and then in year 2 to investigate the effects of external and internal factors on L2 development. Children spoke Singaporean Mandarin and English and completed tests on 3 English language constructs: phonological awareness, receptive vocabulary, and word reading. Results showed that, in general, internal factors had the largest impact on those constructs. However, frequency (per day) of reading English books at home — which was an external factor — loaded as the strongest factor for English word reading skill in the second year of kindergarten in comparison with the previous year (Sun; Bornstein; Esposito, 2021, p. 1760). This detail is not discussed by Sun, Bornstein and Esposito, but is evidence for many hypotheses on how to improve reading skills.

Hassan *et al*. (2021) looked into the influence of reading habits on reading achievement in another language. They had 351 secondary students who were learning a foreign language to participate in this study. Participants answered a linguistic background questionnaire and a standard reading skill assessment. Results showed that reading habits were positively correlated with reading achievement in a foreign language. From Rudell and Hu (2010) to Hassan *et al.* (2021), the effects of reading habits were seen within one language. There are a few pieces of evidence suggesting there might be influences across languages as well.

Camiciottoli (2001) was interested in the frequency of reading in a foreign language by Italian speakers. A questionnaire was administered to 182 Italian-English university students and analyses showed that L1 reading frequency correlated positively with L2 reading frequency. She explains that participants reported low-frequency of reading in their L1 and contemplates that the students' reading habits might be carried over to the other language. Consequently, the one who read less frequently in the L1 will also read less frequently in the L2.

Artieda (2007) examined the influence of L1 literacy and reading habits in linguistic development of adult bilinguals with low or intermediate L2 proficiency. Results demonstrated a moderate correlation between L2 development and L1 reading comprehension for beginner and intermediate learners of English. L1 spelling had an impact on L2 development on both beginner and intermediate learners, but weakly for the latter group. Finally, only intermediate learners were influenced by reading habits. Thus, it is possible to see that L1 literacy may work as a threshold for learning a L2 and that L1 spelling skills are important for developing the language even in intermediate learners. It is suggested that more advanced learners may be more benefited in their L2 development by L2 reading habits than L1 spelling alone.

Santos-Díaz (2017) investigated the impact of reading experience on L1 (Spanish) and L2 (English or French) vocabularies. In order to measure the available vocabulary, graduate students wrote twenty words related to nine themes of interest in everyday life. Also, they completed a specialized vocabulary test and a reading frequency measure based on the number of books read. Generally, participants who reported reading more books also identified a bigger number of specific technical terms, indicating that vocabulary increased as reading in English became more frequent; however, the same did not occur with French. In addition, there was a stronger correlation between reading frequency and active and passive vocabulary in foreign languages than in the L1. This may be explained by similar educational levels across participants, which might equate participants' L1 vocabulary. These findings suggest that L2 reading experience contributes to vocabulary development.

Nonetheless, Caylak Toplu and Erten (2023) present results that differ from the studies reported above. They aimed at examining the association among vocabulary, reading habits and reading motivation for both languages of Turkish-English university students. Participants were 490 Turkish-English university students who completed 2 vocabulary tests and 4 questionnaires on reading motivation and reading habits. However, for this sample, reading habits had no large impact in other constructs. L1 reading habits did predict L1 vocabulary size, which in turn predicted L2 vocabulary size, but L2 reading habits did not explain L2 vocabulary size.

In sum, reading habits are associated with speed of word recognition, that is, the more experience readers have with a language, the faster they will identify words of that language. This is also true for reading habits from outside a formal instruction setting in that they contribute to better word recognition skills in alphabetic and non-alphabetic languages. More frequent reading habits in L1 and/or L2 are related to larger active and passive vocabulary in these languages and are also positively correlated with reading achievement. In regards to reading habits across languages of a bilingual person, the frequency of reading in one language may influence the frequency of reading in the other language. Also, L2 reading habits seem to be especially more linked to L2 development of advanced learners than of beginner ones. The frequency of reading seems to have more effect on active and passive vocabularies of foreign languages than of the L1. Finally, L1 reading habits might influence L2 vocabulary size, but only indirectly.

Thus, it is safe to state that reading in the L1 contributes to reading skills in the same language and, at least indirectly, in the other language. For this reason, reading habits will be included as an independent variable in Study 2.

The next chapters of this dissertation are organized as follows: Chapter 3 presents methods, results and discussion sections for Study 1; Chapter 4 presents methods, results and discussion sections for Study 2; Chapter 5 presents a general discussion considering both studies; Chapter 6 presents an experience report describing difficulties and more details about the data collection process; Chapter 7 presents final remarks.

3 STUDY 1

A bilingual person has knowledge of two languages or dialects and makes use of them (Grosjean, 2008). This may be extended for the multilingual person and their many languages, but possibly involving more complexity. These languages are not separated into hermetic boxes; they do interact and influence language comprehension and production. This has been extensively shown with words. During spoken or written recognition of words, a bilingual or multilingual person may experience effects from words which share features in both languages. One of these effects is the cognate effect, in which the similarity in form and meaning may facilitate or interfere with language processing. For example, the fact that a word is cognate across Spanish and English can make it easier to recognize or to speak than a word which is not cognate (Dijkstra *et al*., 2010; Lemhöfer *et al*., 2008; Schwartz; Kroll; Diaz, 2007; Thierry; Wu, 2007).

In order to test this effect with a language pair which has not been considered in studies on this topic before — Haitian Creole and Brazilian Portuguese —, I designed two studies: one with children and teenagers as participants, and another with adults as participants. In this Chapter, I will describe Study 1, which investigated this psycholinguistic phenomenon in children and teenagers. Study 1 also considered whether phonological awareness (PA) in BP would influence word recognition and the cognate effect. The impact of PA was included because of a vast number of studies pointing a mutual and positive relationship between PA and learning to read. There are also intervention studies showing this effect, such as the study by Santos and Befi-Lopes (2012) carried out in Brazil.

This chapter is composed of the method used in Study 1, presenting research questions, objectives and hypotheses; information about selection of stimuli for tasks, participants, instruments, procedures, experimental design, and data analyses; the results with descriptive and inferential statistical analyses, and the discussion of these results in relation to the hypotheses for the study.

The Haitian Creole speaking population in Brazil arrived in the country migrating from Haiti and the Dominican Republic due to a sociopolitical crisis which is present in Haiti since before the earthquake in 2010. So, in Study 1, I expected to observe how the knowledge of HC either as a first or a heritage language would influence the word recognition skills of children and teenagers who were developing their linguistic knowledge in school in BP as a majority language. In Study 1, there were 33 female participants and 15 male participants ($N = 48$) whose mean age was 12.56 (age range = 8 to 17, $SD = 3.01$). They were children ($N = 17$, $M_{age} = 9.06$;

SD = 0.80) and teenagers (*N* = 31, M_{age} = 14.48; *SD* = 1.81) speakers of HC and BP and were enrolled in public schools in Brazil. A more detailed description of participants is offered in Section 3.1.5.

Participants completed five tasks and one questionnaire. These tasks were the following ones:

1. a visual lexical decision task in BP, with cognate and noncognate words across HC and BP;
2. an auditory lexical decision task in BP, with cognate and noncognate words across HC and BP;
3. a BP phonological awareness test based on Capovilla and Seabra (2013);
4. a letter identification test in BP based on Capellini, Oliveira, and Cuetos (2014);
5. a HC receptive vocabulary test based on Lima (2007);
6. and a language history and use questionnaire based on Li *et al.* (2019).

The construction of these tasks is described in detail in Section 3.1.4.

After school administration and teachers allowed it, participants' legal representatives had given consent, and students confirmed to be willing to participate in the study, data collection sessions took place in schools where participants were enrolled. The procedures and steps used during data collection sessions are listed in Section 3.1.6. Inferential data analyses examined whether the cognate status of words would have an effect on RTs and accuracy of participants during written and auditory LD tasks and whether PA would have an effect on these variables and/or would influence the cognate effect. A brief summary of the results would be that no significant cognate effect was seen in any task modality or any independent variable. On the other hand, higher PA in BP scores were significantly associated with higher accuracy rates in the WLD task. The results and data analyses are reported in detail in Section 3.2.

3.1 METHOD

3.1.1 Research questions, specific objectives, and hypotheses

Study 1 investigated whether children and teenagers who are native speakers of HC show effects of cross-linguistic interaction between HC and BP on lexical access during comprehension of spoken and written BP. The study was guided by the following research questions:

1. Do cognate words across HC and BP, in comparison with noncognate words, facilitate visual and auditory lexical access in BP?
2. Does phonological awareness (PA) in BP influence visual and auditory lexical access in BP?

In order to answer these research questions, Study 1 pursued the following specific objectives:

1. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during visual and auditory lexical decision tasks in BP;
2. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce higher accuracy rates during visual and auditory lexical decision tasks in BP;
3. To investigate whether higher PA scores in BP, in comparison with lower PA scores, produce shorter RTs in visual and auditory lexical decision tasks in BP;
4. To investigate whether higher PA scores in BP, in comparison with lower PA scores, produce higher accuracy rates in visual and auditory lexical decision tasks in BP.

Objectives 1 and 2 are related to research question 1, while Objectives 3 and 4, to research question 2. The Hypotheses for Study 1 are:

1. Cognate words, in comparison with noncognate words, produce shorter RTs during both visual and auditory lexical decision tasks in BP.
2. Cognate words, in comparison with noncognate words, produce higher accuracy rates during both visual and auditory lexical decision tasks in BP.
3. Higher scores in PA in BP, in comparison with lower PA scores, produce shorter RTs in both visual and auditory lexical decision tasks in BP.
4. Higher scores in PA in BP, in comparison with lower PA scores, produce higher accuracy scores in both visual and auditory lexical decision tasks in BP.

The rationale for hypotheses 1 and 2 follows from the evidence that there are variables which impact word reading (Yap; Balota, 2015). Two of those variables are orthographic and phonological similarity. Kroll and Ma (2018) mention that there is strong evidence pointing to the influence of words which share features across languages, such as cognates and interlingual homographs or homophones. In the case of cognate words, orthographic, semantic and, sometimes, phonological representations overlap between languages, accumulating activation. It has been shown that this coactivation may cause either interference or facilitation in language processing. For instance, Dijkstra *et al*. (2010) showed that, in a lexical decision task, cognate words produced inhibition effects. Similarly, Lemhöfer, Huestegge, and Mulder (2018) observed that bilinguals had more difficulties in rejecting near-cognate words in a visual lexical decision task. However, in the study by Lemhöfer *et al.* (2008), results showed that, between languages, cognate status was the strongest predictor of L2 word recognition and that bilinguals benefited from the presence of cognate words. In addition, Mulder *et al.* (2022) observed that, during an auditory lexical decision task, bilingual people responded faster to cognate words in comparison with noncognates. This cognate facilitation was present only when stimuli were the full form of the word, not a reduced form, which was also manipulated in the study. However, higher proficiency levels seem to be associated with smaller cognate effects (Titone *et al.*, 2011). Bultena, Dijkstra, and van Hell (2014) showed that cognate facilitation effects occur in L2 sentence context, but they were smaller for participants who reported high L2 proficiency. In the present study, participants were not highly proficient in BP. Taken together, the evidence mentioned above justify expecting that cognate words may lead to facilitation effects in the visual and auditory lexical decision tasks used in the present study.

The rationale for hypotheses 3 and 4 is inspired by studies which show that phonological awareness and word reading are mutually related (Wagner; Piasta; Torgesen, 2006). Considering these as skills, many abilities seem to be related to reading and its learning, such as spelling and listening comprehension (Verhoeven; van Leeuwe, 2012). When examining the effects of an audio-verbal skills stimulation program on phonological awareness, Toffoli and Lamprecht (2008) observed a significant, strong and positive correlation between participating in the program and performance in phoneme-level tasks. This may indicate that speech perception and phonological awareness are also associated.

The construct phonological awareness was chosen instead of phonemic awareness because the literature indicates that phonological awareness seems to start being developed earlier than phonemic awareness (Anthony; Francis, 2004, 2005; Anthony *et al*., 2003; Milankov *et al*., 2021). Although phonemic awareness seems to have a stronger relationship with word decoding (Mann; Foy, 2003) than phonological awareness, the latter appears to influence reading before schooling starts (Anthony *et al*., 2003; Carroll *et al*., 2003) at least in more transparent orthographies (Milankov *et al*., 2021).

Further, this scenario could be extended to learning to read in a second language. Saiegh Haddad (2019) explains that there is phonological awareness specific for an L2 and that it is dependent on accurate L2 lexical representations. The author adds that oral L2 linguistic knowledge is important to develop L2 phonological awareness, which is also influenced by the phonological structure of the languages. This emphasizes the idea that phonological awareness is language-specific and is associated with different linguistic skills. In addition, transfer of some skills from L1 to L2 has been already reported. Verhoeven (2007) explains that the development of L1 and L2 of minority children is interrelated and also associated with phonological awareness. The author adds that L1 and L2 skills predict independently phonological and phonotactic awareness. This indicates that, if consistently exposed to bilingual educational contexts, children are able to transfer some skills from one language in order to improve learning of the other one.

3.1.2 Experimental design

Study 1 is a cross-sectional, quasi-experimental study (Christensen; Johnson; Turner, 2015). The independent variable is the cognate status of words, either cognate or noncognate, which means that the variable is categorical and has two levels. The dependent variables are response times (RTs) in milliseconds, which are a continuous variable, and response accuracy, which is a binary variable, in both the visual and the auditory lexical decision tasks. The covariable is BP phonological awareness, a continuous variable measured by the scores in the phonological awareness test.

3.1.3 Selection of cognate words

Study 1 manipulates language co-activation by including cognate words across HC and BP in the stimuli list. Two lexical decision tasks were constructed, a visual one and an auditory one, which had unique lists of cognate and noncognate words. This section describes the process and the steps taken to select cognate and noncognate words for the lexical decision tasks.

Ideally, for any verbal task in a Psycholinguistic study, words should be selected from an annotated corpus so the characteristics of these stimuli words come from natural texts (Reppen; Simpson-Vlach, 2019, p. 92) and can be controlled for. However, there were 3 problems regarding the use of a HC corpus in the present study. During the process of word selection, I found a Haitian Creole Wikipedia corpus based on material from 2021, in the Leipzig Corpora Collection (Leipzig, 2024). This dataset has 31,400 sentences and 433,294 tokens in HC. However, the files containing words was not annotated nor normalized. These were text files with the character or word and a number for frequency value. This means that it was not possible to distinguish words according to parts of speech or to know the frequency proportion of each word in the corpus. HC has particles as grammatical structures, and these particles should be avoided in the experimental stimuli lists. Since all stimuli should be controlled to be highly frequent, information about frequency values and frequency proportion in the corpus was necessary. In addition, it was not possible to compare different values of word frequency since they were not normalized in the corpus. There were frequency values, but there was no proportion to compare the values to. Finally, the corpus I had available did not have information about the frequency of occurrence of HC words used as a minority language setting in the context of migration.

Thus, I constructed a corpus-like set of HC words and their translations into BP with the help of an informant. The male informant came from Haiti to Brazil in September 2014 and was highly educated. The initial list of words was extracted from a Haitian Creole-English dictionary (Targète; Urciolo, 1993) because it presented information about parts of speech. The dictionary compiles approximately 60 entries per page and has a total of 236 pages of entries, which equals approximately 14,160 entries. I extracted all nouns, adjectives (which are called attributives in the dictionary) and verbs and organized them in one spreadsheet.

With colleagues from the Laboratório da Linguagem e Processos Cognitivos (LabLing, Language and Cognitive Processes Laboratory) at UFSC, I used the English translations from the HC-Eng dictionary as support for translating the words into BP. The informant reviewed these translations to make sure that they would keep the meaning of the original HC word as best as possible and that they would be commonly used. This means that the BP translations of the HC words are biased by the informant's knowledge of and familiarity with the HC language, which is a limitation of the present study. In order to compensate for this bias, I contacted another informant and invited this person to, voluntarily, help reviewing the translations. However, they could not aid due to a full work schedule. Therefore, the translations were reviewed by only one informant.

In parallel to the dictionary, a glossary was used. Bon Bagay HC-BP and BP-HC glossary (Cotinguiba; Cotinguiba; Andretta, 2018) was created in Porto Velho, Rondônia. It offers entries for more than 5,200 words in HC. The Cotinguiba couple started the glossary as an extension project in 2011 to help a community of Haitian immigrants in Porto Velho. The project was expanded and received help from groups from many universities — such as Federal University of Rondônia, Valparaiso University, Unicamp, Ufam, Unifap, UTFPR, Unochapecó, UFRGS, and University of Brasília — and from CAPES as a funding agency. There was also support from a linguist, socio-anthropologist and educator, Professor Obrillant Damus (State University of Haiti and Quisqueya University). Since the glossary did not include information on parts of speech, it was used to solve translation problems.

Other than reviewing the translations, the informant was asked to classify every word according to frequency of use. The scale was as follows: 0 equalled “never used”, 1 equalled “seldom used”, 2 equalled “frequently used”, and 3 equalled “always used”. These criteria are also considered limitations because they are subjective measures from only one informant. In addition, the researcher and colleagues did not add to the set of potentially experimental words any strongly negative words, such as curse words and words related to death, suffering, sexual intercourse, sexual body parts or innuendos.

Cognate status of words across languages was calculated using the orthographic similarity calculator available at NIM (Guasch *et al.*, 2013), a search engine for psycholinguistic research tools. The calculator uses a formula adapted by van Orden (1987) from Weber (1970). All words and their translations were entered into the calculator. Word pair which presented an orthographic similarity score of 0.600 or higher were considered cognates, and 0.399 or lower were considered noncognates. Experimental stimuli were not selected among word pairs which scored between 0.401 and 0.599. There were 546 high-frequency words categorized as cognate, 3,578 high-frequency words categorized as noncognate, and 1,242 high-frequency words categorized as neither cognate nor noncognate. Examples are presented in Table 2. Cognate status calculations were based on only orthographic similarity because phoneme-grapheme conversion rules used to write HC are very consistent (Breiter, 2014) and because different accent pronunciations would not be considered distinctive when categorizing cognate and noncognate words.

Finally, the entire set of HC words translated into BP, with annotations for parts of speech, frequency of use and cognate status, had 5,617 entries. It is available on OSF (https://osf.io/mcvf8). When selecting words to be used in the experimental tasks, only “always used", that is, high-frequency words were chosen. Table 2 presents examples of the high frequency words.

Table 2 – Examples of high-frequency items in the final HC-BP word set

Word in HC	Word in BP	Part of speech	Frequency of use	Orthographic similarity	Cognate status
minit	minuto	noun	always	0,611	cognate
problèm	problema	noun	always	0,637	cognate
krent	medo	noun	always	0,064	noncognate
odas	audácia	noun	always	0,063	noncognate
depotwa	depósito	noun	always	0,534	left out
vizitè	visitante	noun	always	0,441	left out

Source: The Author (2024)

3.1.4 Instruments

All tasks in Study 1 can be accessed through the following link: https://antigo-labling.ufsc.br/estudos/pietra/criancas/. The final stimuli list for each task is available on OSF (https://osf.io/c5pzj/files/osfstorage). These tasks are a visual and an auditory lexical decision task in BP, a PA task in BP, a letter identification task, and a HC vocabulary task. Also 20 questions from the Language History Questionnaire (LHQ 3.0) were selected to be answered by participants or their families. These tasks are described below.

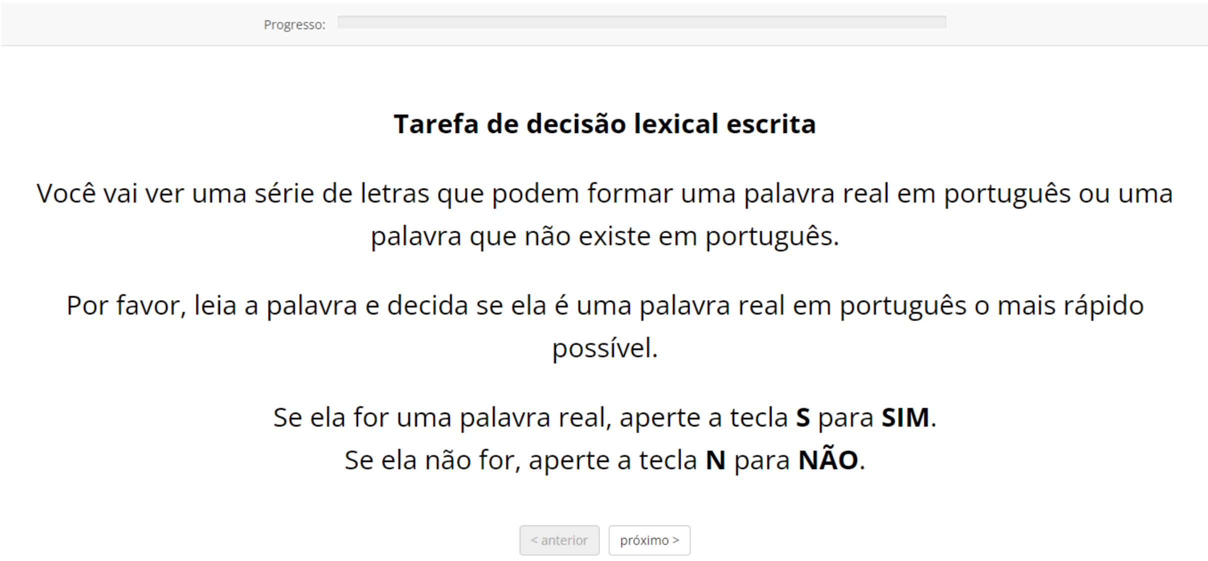
3.1.4.1 The written lexical decision (WLD) task

The written lexical decision task had a traditional visual lexical decision design, that is, participants are required to decide if a string of letters is a real word in one of the languages they use — in this case a real word in BP —, compared to a generalized lexical decision task, in which participants decide if a string of letters is a real word in any of the languages they use (Dijkstra, 2005). I named it *written* instead of *visual* in order to more easily remember that the stimuli were presented in written form on the computer screen as opposed to being auditorily presented. The stimuli consisted of 120 real words and 120 pseudowords in BP. The 120 real words consisted of 60 cognate words across HC and BP (for example, *memória*, *viver* and *mal*) and 60 noncognate words, that is, simply real words in BP (for example, *carro*, *carregar* and *nervoso*). Out of the 120 pseudowords, 72 (for example, *árovre*, *alpica* and *ideisa*) were taken from Justi, Justi & Rossi (2014). Since all of their stimuli were 6-letters long, another 48 pseudowords were created to ensure that the variation in the length of pseudowords was similar to the variation of the real words. These pseudowords were created from cognate and noncognate words which were not used in the lexical decision tasks. Up to two modifications were made in each of those words, either replacing a letter with a new one or switching two existing letters (for example, *uçar* from *usar*, *afual* from *atual* and *irratinte* from *irritante*). It was not possible to control for stimuli length and have all stimuli be at constant length because there were not enough high-frequency short cognate words to be used in the two lexical decision tasks (auditory and visual) and, in the case of Study 2, in both sentence comprehension tasks (auditory and visual). In addition, there were not enough unique high-frequency cognate nouns across HC and BP to be used in all tasks. Thus, a methodological decision was taken to control for frequency and to add different word lengths and word classes in the tasks. Cognate words length ranged from 3 to 12 letters (M = 6.67; SD = 1.74). The length of noncognates ranged from 3 to 12 letters (M = 6.70, SD = 1.98). For cognate words, there were 30 nouns, 15 verbs and 15 adjectives, and for noncognates there were also 30 nouns, 15 verbs and 15 adjectives.

The task was programmed in JavaScript using the jsPsych framework (de Leeuw; Gilbert; Luchterhandt, 2023). The entire task was in BP. The only way we would observe HC co-activation would be if cognate words were processed faster or slower than noncognate ones. First, participants read the instructions shown in the center of the screen while the researcher presented the same instructions orally. Then they clicked on “Next” to start a practice session with the visual stimuli. There were 2 cognate words, 2 noncognate words, and 3 pseudowords in practice trials, and there were 240 experimental trials in total. Instructions were repeated orally in a more concise way before the actual task began to make sure instructions were fresh in the participants’ memory. If participants completed the practice trials correctly, they were allowed to start the task; if they made mistakes, they were motivated to repeat the practice trials.

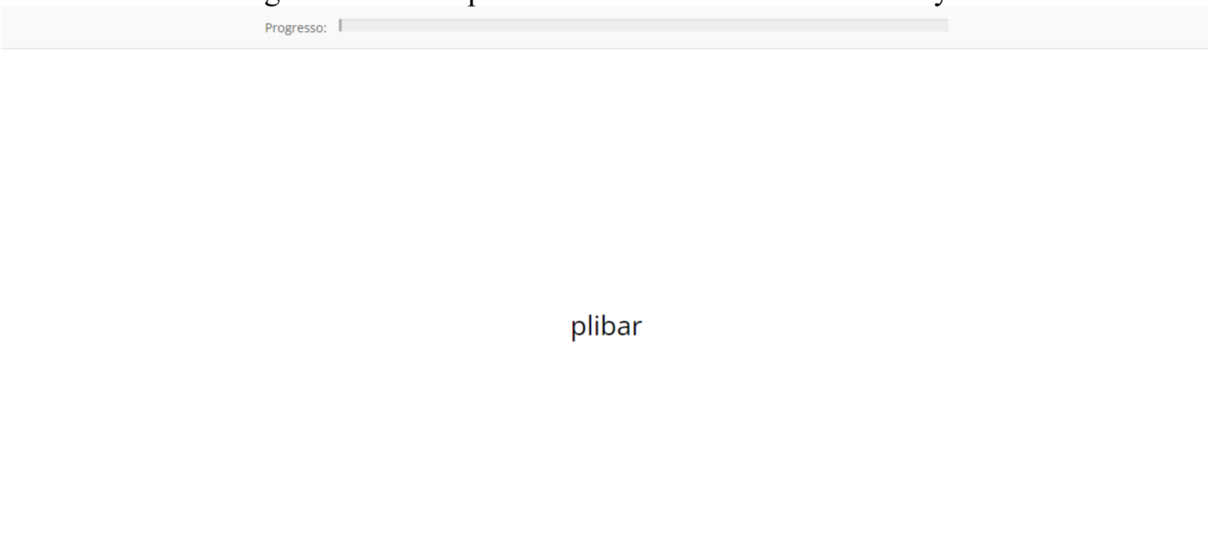
All stimuli were presented in the center of the screen, in lowercase, using font Arial Sans Serif, size 24, in black. Stimuli were displayed on screen until the participant gave an answer. Before each trial, a fixation cross was shown on the center of the screen for 2 ms. The background was white, and there was a gray progress bar at the top. Stimuli were presented in randomized order. Laptop and computer specifications are presented in the Procedures section (3.1.6). Figures 11 and 12 show the instructions screen and an example trial screen of the WLD task. The final stimuli list used in this task can be examined directly in OSF (https://osf.io/3qxdm).

Figure 11 – Instructions screen for the WLD task in Study 1



Source: The Author (2024)

Figure 12 – Example of trial for the WLD task in Study 1



Source: The Author (2024)

Response times and accuracy were registered from the moment the stimulus started to be presented. Stimuli were kept on the screen until the participant responded. During trials, the keyboard only accepted input from the keys N and S. Participants would press N when the stimulus was not a real word in BP and S when it was indeed a real word in BP. They were instructed to answer as fast and accurately as they could. Participants took an average of 10.13 minutes (*SD* = 5.09) to complete the written lexical decision task. There were 7 practice trials and 240 experimental trials in total. Details about participants and procedures will be given in sections 3.1.5 and 3.1.6.

3.1.4.2 The spoken lexical decision (SLD) task

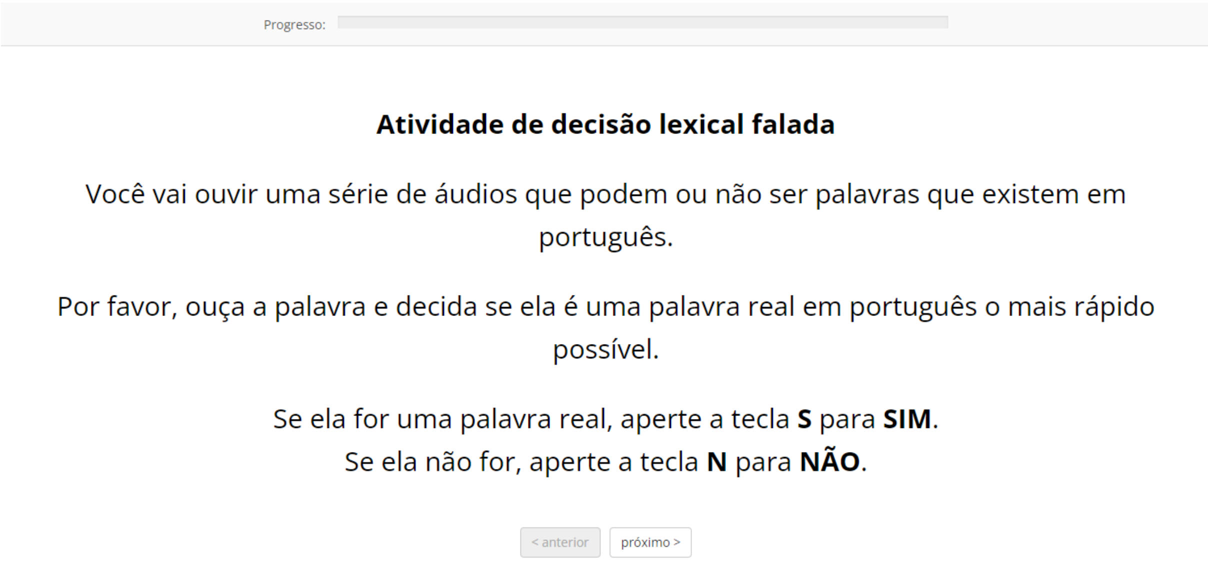
Spoken lexical decision was assessed by means of an auditory lexical decision task. Task structure characteristics and programming were the same as in the written lexical decision task, described in section 3.1.4.1. The major difference between the *written* LD task and the *spoken* LD task was that all stimuli were presented *auditorily* in the spoken LD task.

The selection of the stimuli for the SLD task followed the same criteria as those described in the WLD task. In addition, all words were recorded by two professional voice actors, a man and a woman, contacted via the Vinte Pila platform[1]. This means that there was an audio file recorded with a male voice and with a female voice for every word in the task. In order to counterbalance the female and the male voice across stimuli presentation, the task was programmed so as to select randomly one of the two voices when presenting an item. In other words, some participants may have heard a given word spoken by a male voice and other participants may have heard it spoken by a female voice.

The length of cognate words ranged from 3 to 11 letters (M = 6.72, SD = 1.95). Similarly, the length of noncognates ranged from 3 to 11 letters as well (M = 6.65, SD = 1.90). First, participants read the instructions shown in the center of the screen while the researcher presented the same instructions orally (Figure 13). Participants clicked on "Next" to start a practice session with the auditory stimuli. The practice session consisted of 2 cognate words, 2 noncognate words, and 3 pseudowords, that is, 7 trials in total. If participants completed the practice trials correctly, they were allowed to start the task; if they made mistakes, they were motivated to repeat the practice trials. After finishing the practice session, instructions were repeated orally in a more concise way before the experimental session began.

Figure 13 – Instructions screen for the SLD task in Study 1

[1] Vinte Pila is a Brazilian platform which allows people to offer a wide variety of services for affordable prices. These services are related to areas such as digital marketing, graphic design, translation, audio and video editing, programming etc. I looked for announcements of voice actors offering to record messages in audio. I contacted two women and one man. Before contracting their service, I asked the voice actors to record a sample of phonemes for the Letter Identification task. Their performance in this sample recording would help me choose the voice actors. The first woman did not understand how to record phonemes, but the second woman and the man understood my request and executed it according to my instructions. All auditory stimuli were recorded by these two voice actors. I explained why I needed the words, the pseudowords and the phonemes to be recorded and I instructed the woman and the man how to pronounce the stimuli so I could edit the audio file. I paid for their service.



Source: The Author (2024)

The experimental session of the SLD task consisted of 240 experimental trials in total. Both in the practice session and in the experimental session, a fixation cross was shown on the center of the screen for 2 ms before each trial. Participants were instructed to listen to the audio and to press S if the sequence of sounds was a real word in BP or N if it was not a real word in BP as fast and accurately as possible. I informed participants orally that it was not possible to replay a word so they had to pay attention to the audio. The next stimulus item only played after participants provided an answer using the keyboard.

Participants took an average of 9.83 minutes (*SD* = 2.23) to complete the spoken lexical decision task. During experimental trials in the SLD task, only the progress bar of the task was shown in the screen while the audio stimuli was playing.

The stimuli list and the audios used in this task can be examined directly in OSF (respectively, https://osf.io/5qdpe, https://osf.io/dfrts, and https://osf.io/zeckq).

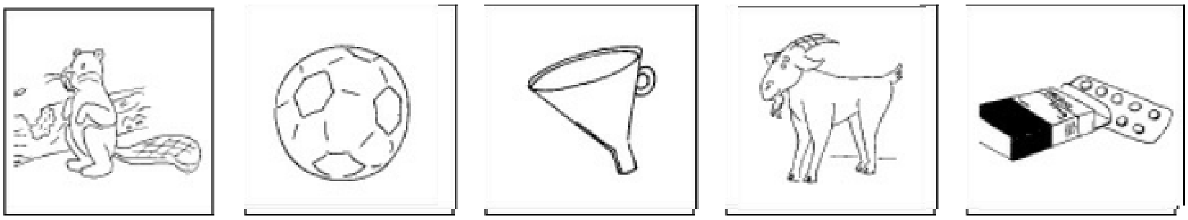
3.1.4.3 The BP PA test

The BP PA test consisted of part of the Prova de Consciência Fonológica por Escolha de Figuras (Picture Selection Phonological Awareness Test) by Capovilla and Seabra (2013), published in Seabra and Dias (2013, chapter 6). It contains nine subtests with two practice trials and five test trials each. The Picture Selection Phonological Awareness Test was designed for children in elementary school and requires them to point to one of five pictures in order to answer the question spoken by the evaluator. Results in this test were shown to be highly correlated with results in the Prova de Consciência Fonológica por Produção Oral (Oral

Production Phonological Awareness Test) according to “$r = 0{,}76$; $r^2 = 0{,}58$; $F(1, 55) = 75{,}14$; $p < 0{,}001$” (Seabra, 2013, p. 144).

In the Picture Selection Phonological Awareness Test, the experimenter is expected to read the questions out loud for the participants. Five pictures are shown side by side as answer options. Participants have to point to the picture which best matches the answer. Response times are not registered, only accuracy. Thus, participants are not required to say the words out loud. For example, the question “What drawing ends with the same word part as the word GLUE?” (in BP, “Qual desenho termina com o mesmo pedaço que a palavra COLA?”) had the following options of pictures:

Figure 14 – Examples of response options in a trial in the PA test in Study 1



Source: Capovilla and Seabra (2013, p. 180)

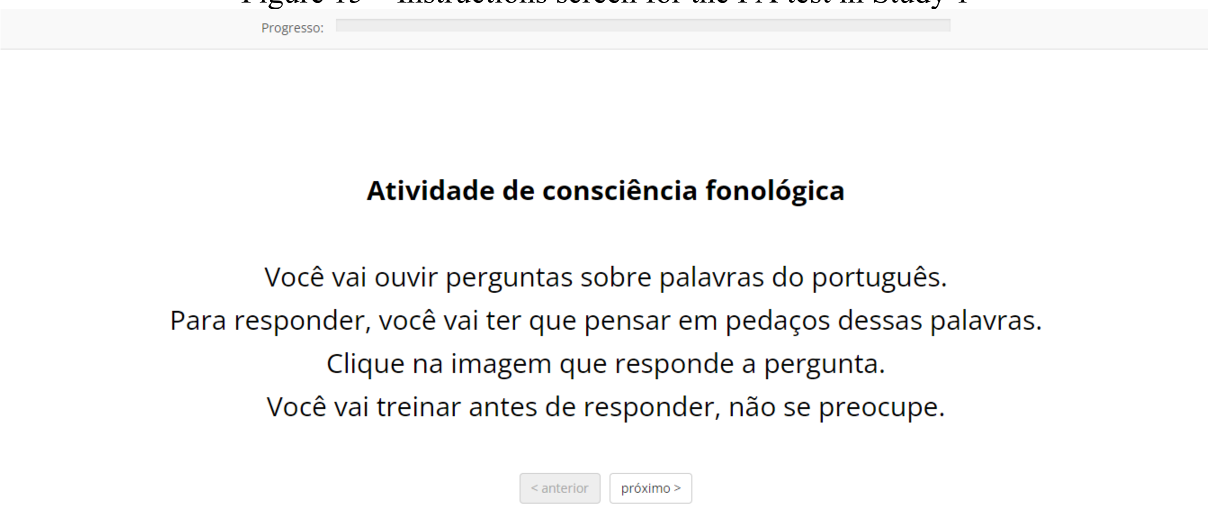
The correct answer here would be “ball” (in BP, “bola”), the second picture from the left.

The Picture Selection Phonological Awareness Test was chosen for the present study for two reasons. The first one is that it requires participants to select pictures in order to answer questions about phonological awareness instead of orally producing the answers. Pointing to a picture had advantages over orally producing words because it was not possible to guarantee that data collection would occur in quiet rooms, which could affect negatively the quality of answer recording. The second reason for selecting this task was that the construction and selection of instruments for the present study took place during 2021. The COVID-19 pandemic was still a global public health emergency during the year of 2022 (CDC, 2023), and most schools were operating remotely in Brazil. To assess PA, I selected a task which did not require spoken responses from participants because, at the time, it was not possible to know whether they would have an ideal environment to perform the tasks, that is, a quiet place, either at home or at school, with good internet connection and a good microphone. However, I was able to start the data collection sessions when schools were already open. Thus, participation sessions occurred in-person, but the PA test was still presented on the web browser with recorded questions in order to keep stimuli questions as controlled as possible.

For Study 1, I adapted the Picture Selection Phonological Awareness Test in case it had to be completed remotely. The only difference between the original test and my adapted version of it was that questions were recorded. In my adapted PA task, the participant would listen to the question and click on the picture which best answered it on the computer screen. This was implemented in the task by playing only once an audio of the question and letting the participant click on the chosen picture on the computer screen among the five answer options. In comparison with the original task, questions are still spoken to the participants, but the answer is given to a recorded question, not to a question asked in person. The wording of instructions and of questions was adapted in a more concise way from the original instructions, which offered the possibility of dialogue between experimenter and participants. Thus, in the PA task in Study 1, no dialogue was expected to happen between participants and experimenter.

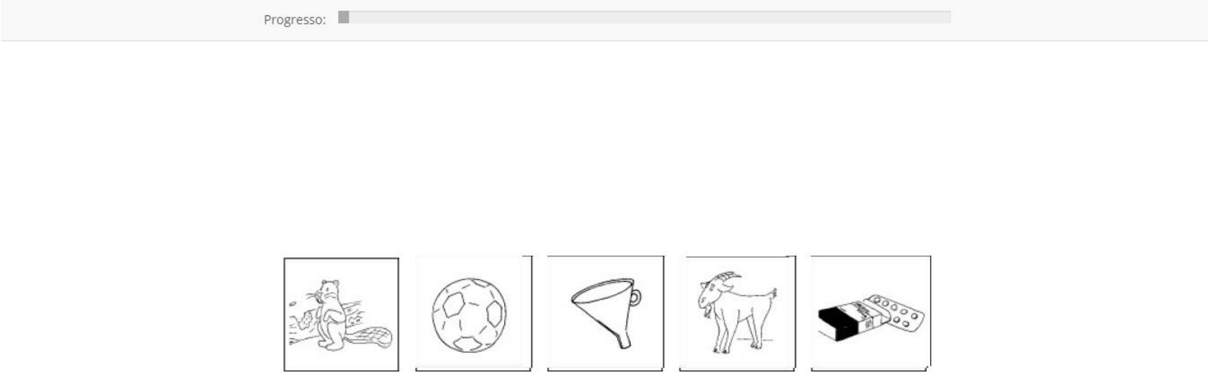
For the PA test in Study 1, I selected 40 experimental questions from the original task, each with 5 options of answers, because the pictures for the 5 last questions were not available. Questions were presented in a fixed block order as follows: 2 practice questions followed by 5 experimental questions. Blocks were not randomized across participants due to programming limitations. There was no time limit for answering a question, that is, participants could take as long as they needed to select the answer. The next question would only be presented after the participant had clicked on a picture to answer a question. Response times were not registered, only accuracy was. Figures 15 and 16 show the instructions screen and an experimental trial screen in the PA task.

Figure 15 – Instructions screen for the PA test in Study 1



Source: The Author (2024)

Figure 16 – Example of the screen of a trial in the PA test in Study 1



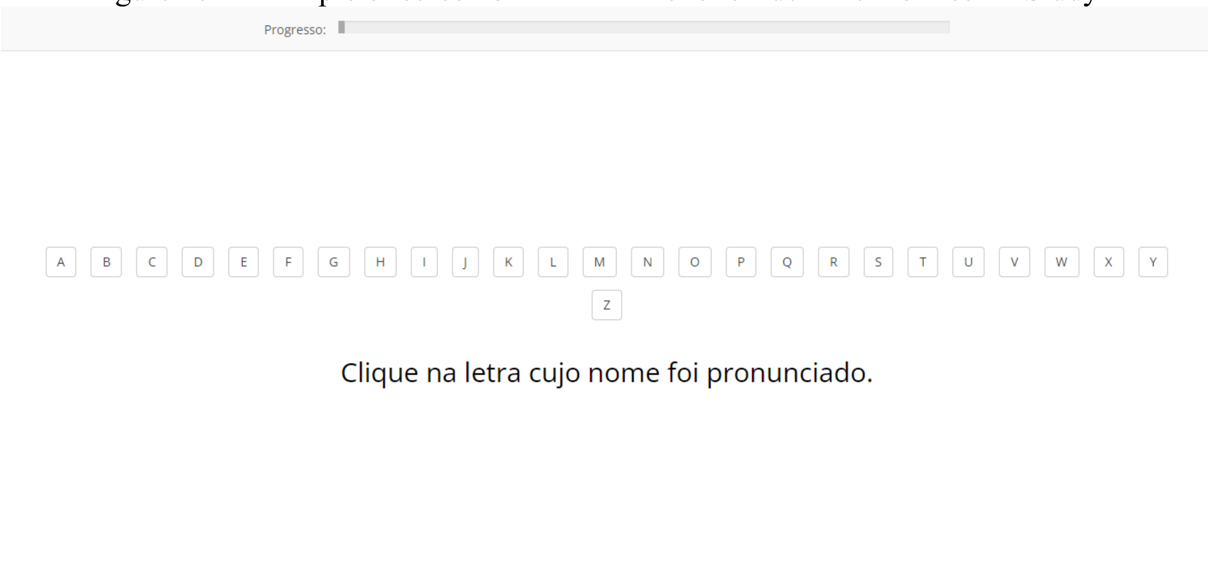
Source: The Author (2024)

All questions and instructions were recorded by the two voice actors who also recorded the audio files for the SLD task. Thus, all questions and instructions were recorded in a female voice and in a male voice. Order of presentation of questions was fixed, as mentioned above, but the voice in which they were recorded was randomized, that is, some participants listened to a question spoken in a female voice while other participants listened to this same question spoken in a male voice. Participants took an average of 12.78 minutes (*SD* = 3.51) to complete this task. The stimuli list, images, and audios used in this task can be examined directly in OSF (respectively, https://osf.io/2fsvn, https://osf.io/hv45x, https://osf.io/qst5a and https://osf.io/6ryhc).

3.1.4.4 Letter identification test

The letter identification test used in Study 1 was added to make sure the children, that is, younger than 11 years old, knew the basics of decoding words in BP. When they scored lower than 50%, their data was not added to analyses. The test was adapted from Capellini, Oliveira, and Cuetos (2014) and, just like in Oliveira, Germano, and Capellini (2016), divided into two parts. In the first part of the test, the participant was required to hear the name of a letter of the alphabet in BP and to point and click on the respective letter written on the screen. Figure 17 illustrates how this was displayed on screen. This was repeated in randomized order through the entire alphabet. In the second part of the test, the participant was required to hear a BP sound (phone or phoneme) associated with a letter of the alphabet and to point and click on any of the possible letters written on the screen associated with that sound. This was repeated in randomized order through the entire alphabet too. There was no time limit for responses, and there was no possibility to hear the audio again. Participants took an average of 4.72 minutes (*SD* = 1.51) to complete this test. The final stimuli list and audios used in this task can be examined directly in OSF (respectively, https://osf.io/jtkyg, https://osf.io/hwdzc and https://osf.io/yv7zc).

Figure 17 – Example of screen of a trial in the letter identification test in Study 1



Source: The Author (2024)

3.1.4.5 HC receptive vocabulary test

This was the only task explicitly tapping HC. Due to the fact that I did not find a validated HC receptive vocabulary test, I adapted the BP version of the Peabody Picture Vocabulary Test created by Lima (2007). She explains that the PPVT is indicated to evaluate either written or auditory receptive vocabulary of people from 2 years and 6 months of age to 18 years of age (Lima, 2007). The test consists of 130 items from varied semantic categories, such as animals, objects, places, body parts, time and nature. The native speaker of HC who served as informant translated all items into HC. If a word had more than one possible translation, the informant was asked to choose the most frequent word form and the most similar meaning as possible.

However, as already explained, I did not know if participants would be able to complete the tasks in a quiet room with a microphone of decent quality. Thus, in order to avoid having to record participants’ answers, I opted to adapt the vocabulary test from a production task to a recognition one. Thus, instead of showing pictures to participants and asking them to name these pictures, the test presented the items in audio format and participants had to answer whether they knew the word or not. More specifically, participants were instructed to listen to a word and to press S if they knew that word or N if they did not know that word. There was no time limit for responding to each item. After providing an answer, the next trial began automatically.

There was a total of 130 trials. Participants took an average of 7.89 minutes (*SD* = 4.47) to complete this task. A fixation cross appeared on screen for 200 ms before the audio stimuli started playing. While the audio for each word was playing, there was no information on the screen except for the progress bar on top of the page. The final stimuli list and the audios for this task can be examined in OSF (respectively, https://osf.io/bxf2k and https://osf.io/sm6f5).

3.1.4.6 Language history questionnaire

The Language History Questionnaire 3.0 (LHQ) was developed by Li *et al.* (2019) and currently has proofread versions in 20 different languages, including Portuguese. It consists of 27 questions. I have used the LHQ in previous research and have noticed that participants tended to take very long to answer all questions in the questionnaire, which makes them tired and impatient. In order to avoid this, I selected 20 questions from the LHQ for this dissertation study. The selected questions inquired about age, education, knowledge of languages, age of onset of learning languages, self-reported linguistic abilities, self-reported proficiency, method of learning languages, language use in different contexts and activities. The selected questions can be examined in Appendix A and their HC versions can be examined in Appendix B. For example, four questions (numbers 10, 14, 15 and 22 in the original questionnaire) would be answered by choosing a number on a 7-point scale. Two other questions (numbers 18 and 19 in the original questionnaire) would require participants to inform the number of hours per day of use per language spoken in specific contexts and for specific activities. Questions were translated into HC by the native speaker of HC so that there was a HC version for participants who would prefer to read an HC document. In Study 1, this questionnaire was completed on paper. Participants who answered the questionnaire themselves took an average of 35.75 minutes (*SD* = 12.19) to complete the questionnaire.

3.1.5 Participants

Initially, participants for Study 1 would be HC-speaking children of 7 to 9 years of age enrolled in public schools in Florianópolis. However, I had to broaden the initial age range and school range because of the difficulty in finding enough children from this group to complete the sample size of 40 participants which we aimed at (Brysbaert; Stevens, 2018). Thus, a decision was taken to invite HC-speaking children and teenagers enrolled in municipal and state public schools in Florianópolis and in state public schools in São José and Palhoça, in the Great Florianópolis area. In Chapter 6, I present a data collection experience report, describing in detail other difficulties encountered during data collection.

A total of 112 schools were contacted via phone calls and e-mails and 30 schools were visited. Fifty-four students were invited to participate in Study 1. The final sample had 48 participants (33 female; age range = 8 to 17, M_{age} = 12.56; *SD* = 3.01), all native speakers of HC.

The criteria for inclusion in the analysis of data were the following:

- No atypical developmental aspect reported by the school;
- Corrected vision or hearing, if impaired;
- Knowledge of at least 50% of the HC words in the vocabulary test.

In order to describe the population sample who participated in Study 1, I will refer to answers given to the 20 questions from the Language History Questionnaire (LHQ) presented in Appendix A. Descriptive statistics about participants' language history and use are presented next. These statistics were only calculated for LHQ items which had answers for at least half of the participants.

Languages spoken, other than HC and BP, were French, Spanish and English. Other than HC, 18 participants reported also speaking French, 6 Spanish, 3 English, 7 French and English, and 2 French, Spanish and English. Mean age of starting to be exposed to HC was 2.16 (*SD* = 1.65) and to BP was 8.96 (*SD* = 3.80). Considering that some participants did not provide answers to some questions, 31 participants reported using BP more at home, 8 reported using HC more at home; 2 reported using HC more at school, 36 reported using BP more at school. Schools helped me with information about the year the 48 students arrived in Brazil. The median year of arrival was 2019 and the mean was 2018,164. Table 3 presents the mean age (*SD* in parenthesis) when participants started using each language in each environment.

Table 3 – Mean age when participants started using each language in each environment in Study 1

(N = 37)	Home	Friends	School	Computer	Online games
Haitian Creole	2.86 (2.20) Range: 1-10	5.00 (2.49) Range: 1-11	4.88 (1.77) Range: 2-9	10.50 (2.72) Range: 6-16	9.75 (2.33) Range: 6-13
Brazilian Portuguese	9.69 (3.49) Range: 3-15	9.37 (3.49) Range: 3-15	9.63 (3.89) Range: 3-17	10.44 (3.01) Range: 5-17	10.00 (3.46) Range: 2-17

Source: The Author (2024)

Table 4 presents means and SDs for self-reported proficiency levels for language skills in HC and in BP. The scale had 7 points, where 7 was the highest proficiency possible.

Table 4 – Mean (SD) self-reported proficiency levels for language skills in HC and in BP in Study 1

(N = 45)	Listening	Speaking	Reading	Writing	All skills
Haitian Creole	5.67 (1.36) Range: 2-7	5.70 (1.34) Range: 1-7	5.70 (1.34) Range: 1-7	4.63 (2.15) Range: 1-7	5.28 (0.44) Range: 4.63-5.70
Brazilian Portuguese	5.96 (1.19) Range: 2-7	6.05 (1.07) Range: 3-7	5.53 (1.38) Range: 3-7	5.22 (1.37) Range: 2-7	5.80 (0.15) Range: 5.53-6.05

Source: The Author (2024)

Table 5 presents means for hours per day using a specific language for each activity;

Table 5 – Mean hours per day using a specific language for each activity in Study 1

(N = 42)	Watching TV	Reading for fun	Reading for school	Writing	Using social media
Haitian Creole	4.00 (4.23) Range: 1-16	1.42 (0.85) Range: 0.5-3	1.80 (1.69) Range: 0.5-5	1.38 (0.65) Range: 0.5-2	4.34 (4.48) Range: 0.5-16
Brazilian Portuguese	5.64 (4.38) Range: 0.5-16	4.12 (5.48) Range: 0.5-20	4.01 (3.97) Range: 0.5-16	4.18 (3.83) Range: 0.5-16	5.79 (5.15) Range: 0.5-20

Source: The Author (2024)

Parents' educational level is reported in Table 6. Some data is missing because some participants did not know their parents` educational level or never sent their answers back.

Table 6 – Number of participants who reported parents' educational level in Study 1

	Preschool	Elementary school	High school	Undergraduate course	Total of answers
Mother	2	17	18	4	41
Father	2	11	21	3	37

Source: The Author (2024)

Study 1 was approved by the Ethics Committee at UFSC under CAAE 51539821.3.0000.0121. To take part in Study 1, participants had to turn in the consent form signed by their legal representative. Participation was only scheduled after the consent form had been returned and indicated the legal representatives' approval. In addition, participants in Study 1 also were presented with the Willingness Form (Termo de Assentimento Livre e Esclarecido, TALE) at the beginning of the data collection session. Participants read the TALE while the researcher read it out loud. Any questions were answered, and data collection only started if participants had agreed to participate. Both the TCLE (Appendices C and D) and the TALE (Appendices E and F) had versions in BP and in HC.

3.1.6 Procedures

Data collection sessions occurred in school rooms according to availability: library, empty classroom, empty coordinator room, empty teachers' room, empty video room, empty multifunction room. Data collection was only scheduled after the TCLE had been returned to the researcher with the legal representative's approval. Participants were welcomed in the room, sat in front of the laptop or desktop computer (which will be described below) and were presented with the website where the experimental tasks were hosted. Tasks were completed either on laptops or desktop computers depending on the machines available at schools. The researcher always took her personal laptop in data collection sessions (DELL Inspiron 3583 15-inch Intel Core i5-8265U, 1.60GHz/1.80 GHz, refresh rate 60Hz). Her research fellow who usually voluntarily assisted during data collection also took her personal laptop (MacBook Air 13-inch 2017, Intel Core i5 1.8 GHz Dual-Core, Intel HD Graphics 6000 1536 MB, refresh rate 60Hz). Desktop computers available in schools were similar to the following model: Intel Core i3-10100T CPU @ 3.00GHz, RAM 8.00GB, 15-inch screen, refresh rate 60Hz. No Bluetooth keyboard was used, only the ones built-into the laptops. It was not possible to control for participants distance from the screen. Headphones mostly used were Philips Bluetooth headphones TAUH202WT/00, 20-20,000 Hz, 102 dB. One third of participants used Sony MDR-EX15LP earphones. Tasks were programmed to be accessible via web browser. In most cases, the web browser used was Google Chrome. However, in the few cases when Chrome was not working, Mozilla Firefox was used.

Data collection instruction for the experimenter are described in the data collection notes file (Appendix K). Participation only began after participants read the Willingness form, asked questions if needed, and agreed to participate. Participants were reminded that they were free to give up participation at any time with no negative consequences for them. The order in which participants would complete the tasks was prepared in advance and counterbalanced as shown below.

1. Spoken lexical decision
2. Written lexical decision
3. Phonological awareness test
4. HC vocabulary test
5. Letter identification task
6. LHQ

Or

1. Written lexical decision
2. Spoken lexical decision
3. HC vocabulary test
4. Phonological awareness test
5. Letter identification task
6. LHQ

If participants were 10 or younger, they would complete the letter identification task first. This decision was a precaution to make sure younger participants knew the BP grapheme-phoneme conversion system. Due to the SARS-COV2 pandemic, Brazilian schools had been reporting that students were showing symptoms of delays in the development of reading abilities. This perceived phenomenon has been recently evidenced by a small-scale cross-sectional study carried out in the city of Sobral, Ceará, in Brasil (Alves; Oliveira; Hirata, 2022). It compared literacy-related scores from two cohorts of students, one from 2019 and the other from 2022. It is possible to observe delays in literacy development comparable to at least 1 year of students' scores. These results confirm that it was correct to expect that not all children under 10 would be fluent in word decoding.

The lexical decision tasks were presented in the beginning of the participation session to make sure participants would not be tired from the tasks. Then, either the vocabulary test was presented or the phonological awareness test. If participants were 10 or older, they would perform the letter identification task at this point. Finally, the LHQ would be completed. Participants younger than 11 would always take the questionnaire home to be answered by family or guardians. Participants 11 or older could answer the questionnaire themselves if they felt comfortable with it. This decision was taken because many families would forget to send the questionnaire back to school or would take longer than 2 weeks to do so. In order to avoid losing the questionnaire answers in this process, pre-teens and teenagers had the option of completing the questionnaire together with the researcher.

Participants were given breaks whenever they needed to use the toilet or to drink water. In those cases, I asked them to finish the task first. Only one participant asked not to wear earphones to complete auditory tasks. Since the room was quiet enough, I allowed it.

3.1.7 Data analysis plan

Data was analyzed using the packages lme4 (Bates *et al.*, 2015) and lmerTest (Kuznetsova; Brockhoff; Christensen, 2017) for mixed-effects linear models in RStudio (Posit Team, 2023). Items and participants were added as random effects. Both random intercepts and random slopes were added at first. If models did not converge, only random intercepts were kept. Four models were run for the planned analyses so that analyses for visual word recognition and spoken word recognition tasks were kept separate.

All planned models had cognate status, phonological awareness, and age added as fixed effects. Models 1 and 2 analyzed data from the WLD task. Model 1 had RTs as the dependent variable for the WLD task, and model 2 had accuracy rates. Models 3 and 4 analyzed data from the SLD task. Model 3 had RTs as the dependent variable for the SLD task, and model 4 had accuracy rates.

Significant interactions were analyzed by running two more mixed-effect linear models, and p-values were corrected by using the function p.adjust() from the stats package (R Core Team, 2023). A pilot analysis was run on the data from 4 participants.

Results section will present first the planned analyses and later the exploratory ones.

3.2 RESULTS

All steps taken during data compilation, cleaning and trimming and during data analysis can be examined in the Rmarkdown file available at the project page for this study on OSF (https://osf.io/df736). First, I will present descriptive statistics and explain the data trimming process. Then, I will present the inferential analyses and the goal for each test. Finally, I will present exploratory analyses which were considered relevant during planned analyses.

3.2.1 Descriptive statistics analyses

Before data trimming, descriptive statistics were calculated. Table 7 shows mean accuracy scores, standard deviations, ranges, and total scores for each task.

Table 7 – Mean accuracy scores, SDs, ranges, and total scores in Study 1

Task	Mean accuracy scores	SDs	Range	Total
Spoken Lexical Decision task	193.646	26.686	131—235	240
Written Lexical Decision task	181.875	29.113	117—230	240
Phonological Awareness test	26.792	7.385	9—38	40
HC Vocabulary test	92.521	25.292	23—126	130
Letter Identification test	40.958	3.808	27—46	52

Source: The Author (2024)

Table 7 also indicates that there were participants who presented error rates higher than 30% in the lexical decision tasks. A score of 168 meant an accuracy rate of 70%. This information was used to select data to include in the inferential analysis. Moreover, scores on the vocabulary test and the letter identification test were used to filter data, which will be described below. Table 8 presents mean response times, standard deviations, and ranges for both lexical decision tasks.

Table 8 – Mean RTs, SDs, and ranges for both LD tasks in Study 1

Task	Mean response times (ms)	SDs	Range
Spoken Lexical Decision task	761.416	2871.546	-1221—218335
Written Lexical	2449.968	11516.810	1—1173031

Decision task			

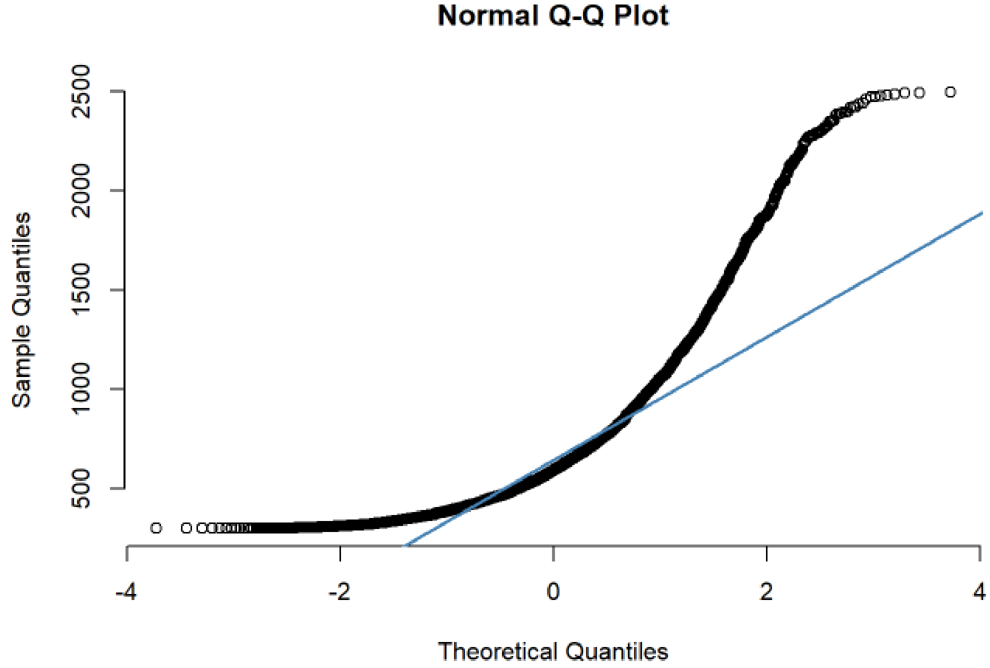
Source: The Author (2024)

Table 8 demonstrates that RTs varied vastly, which may be observed on SDs and range. Considering these characteristics of the data, some trimming was needed. At this point, there were 11,520 observations in each LD task file.

Steps for trimming data were the following. First, I used a threshold of 30% of error rates for removing participant data from further analyses (Arêas da Luz Fontes; Schwartz, 2015). This means that only data from participants who had accuracy scores equal or higher to 70% were included in inferential analyses. At this point, there were 9,600 observations in the SLD task file and 7,200 in the WLD task file. Then, I calculated mean RTs for each participant and removed the ones which were beyond -2.5 SDs or + 2.5SDs from participant mean (Beatty-Martínez *et al.*, 2020). There was not change in number of observations. In addition, I removed RTs shorter than 300 ms and longer than 2,500 ms (Beatty-Martínez *et al.*, 2020). Now there were 6,472 observations in the SLD task file and 5,569 in the WLD task file. Finally, I removed data from participants who reported knowing less than 50% of words in the HC vocabulary test and who had accuracy scores lower than 50% in the letter identification task. At this point, there were 5,055 observations in the SLD task file and 4,989 observations in the WLD task file. This indicates that 56.11% of the data in the SLD task file and 56.69% of data in the WLD task file were removed from further analyses.

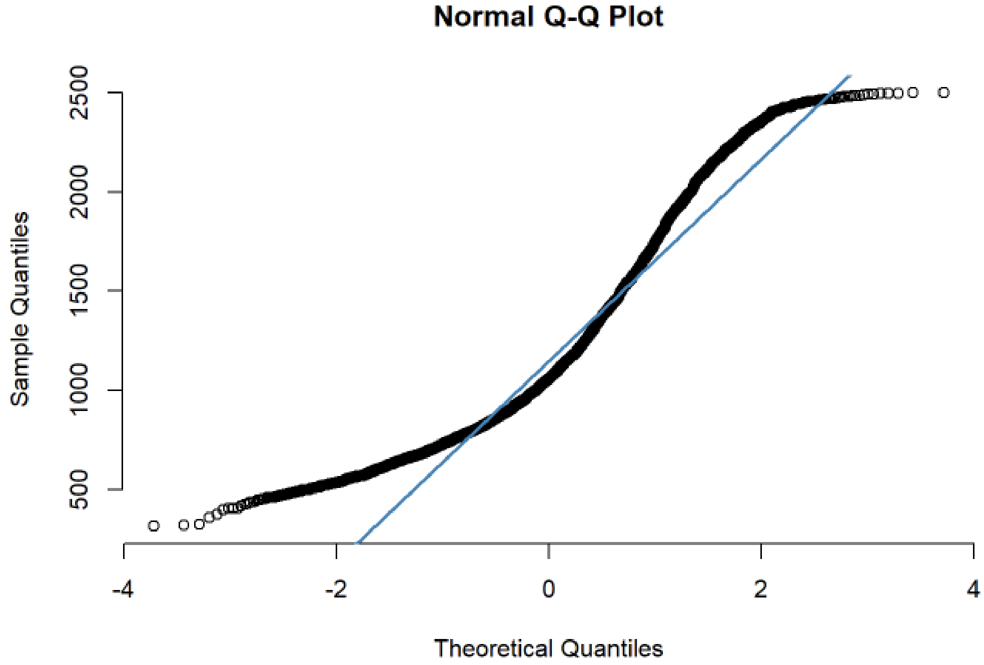
After trimming, the distribution of RTs was examined for normality using the Anderson-Darling normality test (package nortest by Gross and Ligges (2015)) because of the number of observations in the SLD task data file (> 5,000). Response times in both LD task modalities were not normally distributed (WLD task: $A = 120.79$, $p < 0.001$; SLD task: $A = 247.21$, $p < 0.001$). This was also demonstrated with Q-Q plots, which may be examined in Figures 18 and 19, and with histogram plots, which may be observed in Figures 20 and 21.

Figure 18 – Q-Q plot for RTs in the SLD task in Study 1



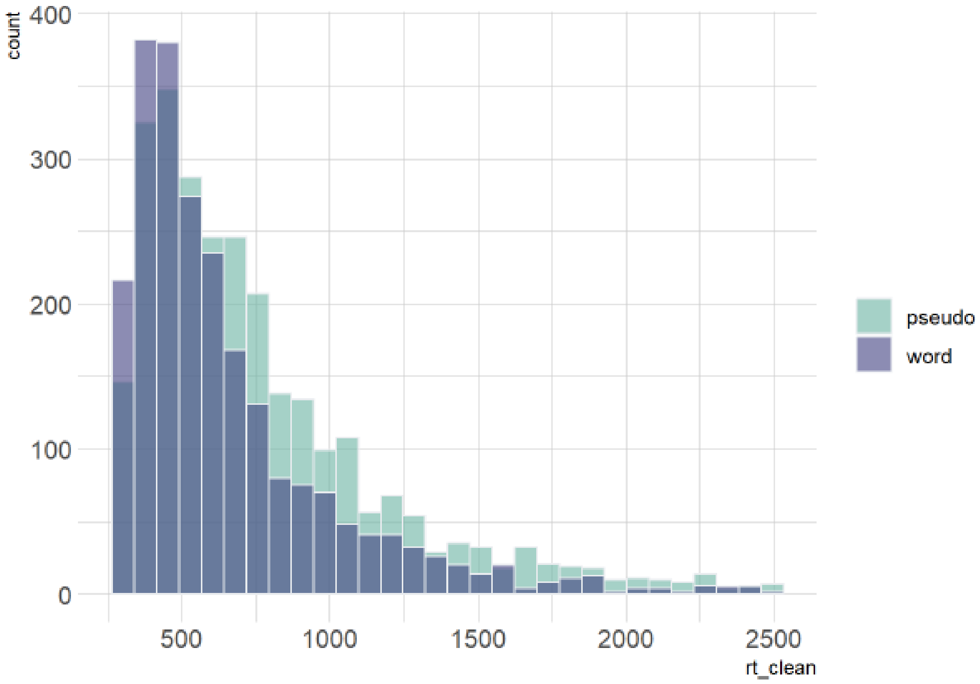
Source: The Author (2024)

Figure 19 – Q-Q plot for RTs in the WLD task in Study 1



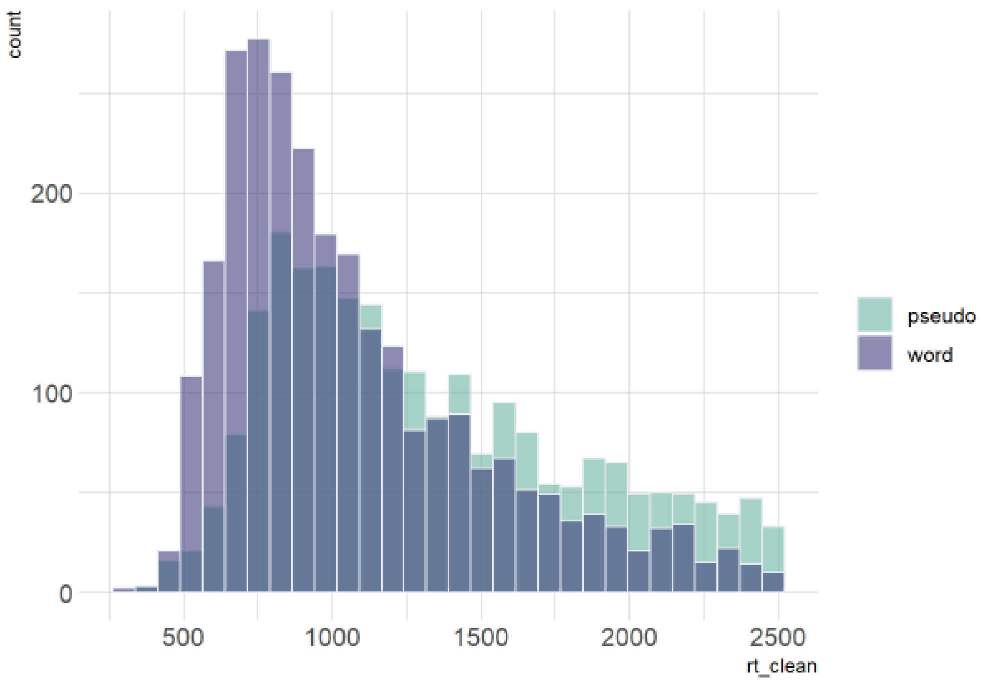
Source: The Author (2024)

Figure 20 – Histogram plot for RTs in the SLD task in Study 1



Source: The Author (2024)

Figure 21 – Histogram plot for RTs in the WLD task in Study 1



Source: The Author (2024)

3.2.2 Inferential statistical analyses

The goal of the lexical decision task is to decide whether a string of letters (or a sequence of phonemes) is a real word or not. One of the effects seen in this task is that real words are responded to faster than pseudowords. Since the distribution of RTs was not normal, I used a Wilcoxon Ranks Sum test to examine whether there was a significant difference in RTs between real word trials and pseudoword trials in both LD task modalities. There was a significant difference between mean RTs for words and for pseudowords in both task modalities (WLD task: $W = 4042029$, $p < 0.001$; SLD task: $W = 3744455$, $p < 0.001$). The descriptive statistics for these differences are presented in Table 9.

Table 9 – Mean RTs and SDs for words and pseudowords in both LD tasks in Study 1

	Trial type	Mean RTs	SDs	Ranges
SLD	Word	657.544	366.522	301-2492
	Pseudoword	764.133	418.025	301-2493
WLS	Word	1075.039	446.381	315-2499
	Pseudoword	1321.335	507.354	324-2497

Source: The Author (2024)

I created four new subsets of the data: two for RT analyses with only answers for words and only correct answers for each LD task modality, and other two for accuracy analyses with only answers for words. I set contrasts and centered all independent variables: cognate status, PA, HC vocabulary, and letter identification (Brehm; Alday, 2022).

Now I will present the planned inferential statistical tests to confirm or to disconfirm the Hypotheses in this study. Model 1 tested Hypotheses 1 and 3 in relation to RTs in the WLD task. Model 2 tested Hypotheses 2 and 4 in relation to accuracy rates in the WLD task. Model 3 tested Hypotheses 1 and 3 in relation to RTs in the SLD task. And model 4 tested Hypotheses 2 and 4 in relation to accuracy rates in the SLD task. I added cognate status and PA together in the models to examine possible interactions. Age was also added as a fixed effect to all models due to the broad age range of participants (7 to 17).

Model 1 had the following code: `m11 <- lmer(data = s1W70_wordcorrect, log(rt_clean) ~ status + PApcent_s + age_s + status*PApcent_s + (1 + status | participant) + (1 | stimulus)).` Table 10 presents estimates, standard errors, and p-values.

Table 10 – Model 1 of planned analyses in Study 1

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.935928	0.054157	25.425420	128.071	<2e-16 ***
Status [cognate]	-0.027020	0.016831	111.835921	-1.605	0.1112
PA (%)	-0.037371	0.049914	21.910578	-0.749	0.4620
Age	-0.136292	0.043392	23.552972	-3.141	0.0045 **
Status [cognate]:PA	-0.002134	0.005940	83.171372	-0.359	0.7203
Random effects					
	Variance		SD		
Stimulus	0.0294459		0.17160		
Participant	0.0653535		0.25564		
Residual	0.0751418		0.27412		
N observations	2466				
N stimulus	120				
N participants	25				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

There was only one significant fixed effect in Model 1: age (corrected p-value = 0.0225). This indicates that the time participants took to decide whether a string of letters was a real word in BP was modulated by their age (β = -0.136292). The older the participants, the faster were their RTs. However, neither cognate status of words or PA in BP had any impact in RTs in this task. There was also no interaction between cognate status and PA. Thus, Hypothesis 1 — that cognate words, in comparison with noncognate words, are associated with shorter RTs during both visual and auditory lexical decision in BP — and 3 — that higher scores in PA in BP are associated with shorter RTs in both visual and auditory lexical decision in BP — were not confirmed.

Model 2 had the following code: `m12 <- glmer(data = s1W70_word, acc ~ status + PApcent_s + age_s + status*PApcent_s + (1 | participant) + (1 | stimulus), family = "binomial")`. Table 11 presents estimates, standard errors, and p-values.

Table 11 – Model 2 of planned analyses in Study 1

Fixed effects					
	Estimates	SE	z value	p	
(Intercept)	4.23899	0.37724	11.237	<2e-16 ***	
Status [cognate]	0.21219	0.25669	0.827	0.4084	
PA (%)	0.57565	0.21114	2.726	0.0064 **	
Age	0.26165	0.19682	1.329	0.1837	
Status [cognate]:PA	-0.07599	0.08238	-0.922	0.3563	
Random effects					
	Variance		SD		
Stimulus	5.0673		2.2511		
Participant	0.8754		0.9356		
N observations	2675				
N stimulus	120				
N participants	25				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

There was only one significant fixed effect in Model 2: PA in BP (corrected p-value = 0.0320). This indicates that the accuracy to decide whether a string of letters was a real word in BP was modulated by the score in the PA test ($\beta = 0.57565$). Participants who scored higher in the PA test also had more correct answers in the WLD task. However, neither cognate status of words or age had any impact in accuracy rates in this task. There was also no interaction between cognate status and PA. Thus, Hypotheses 2 — that cognate words, in comparison with noncognate words, are associated with higher accuracy rates during both visual and auditory lexical decision in BP — was not confirmed, but Hypothesis 4 — that higher scores in PA in BP are associated with higher accuracy scores in both visual and auditory lexical decision in BP — was partially confirmed.

Model 3 had the following code: `m13 <- lmer(data = s1S70_wordcorrect, log(rt_clean) ~ status + PApcent_s + age_s + status*PApcent_s + (1 + status | participant) + (1 | stimulus))`. Table 12 presents estimates, standard errors, and p-values.

Table 12 – Model 3 of planned analyses in Study 1

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.327e+00	2.892e-02	3.601e+01	218.784	<2e-16 ***
Status [cognate]	4.185e-03	1.360e-02	1.020e+02	0.308	0.759
PA (%)	-1.700e-02	2.823e-02	2.844e+01	-0.602	0.552
Age	-3.712e-02	2.728e-02	2.927e+01	-1.360	0.184
Status [cognate]:PA	-1.509e-04	9.239e-03	4.173e+02	-0.016	0.987
Random effects					
	Variance		SD		
Stimulus	0.0114711		0.107103		
Participant	0.0205925		0.143501		
Residual	0.1644434		0.405516		
N observations	2016				
N stimulus	120				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

There were no significant fixed effects and no interactions in Model 3. RTs in the SLD task were not influenced by cognate status, PA in BP or age. Thus, Hypothesis 1 and 3 were not confirmed.

Model 4 had the following code: `m14 <- glmer(data = s1S70_word, acc ~ status + PApcent_s + age_s + status*PApcent_s + (1 | participant) + (1 | stimulus), family = "binomial")`. Table 13 presents estimates, standard errors, and p-values.

Table 13 – Model 4 of planned analyses in Study 1

<table>
<tr><th colspan="6">Fixed effects</th></tr>
<tr><td></td><td>Estimates</td><td>SE</td><td colspan="2">z value</td><td>p</td></tr>
<tr><td>(Intercept)</td><td>2.88133</td><td>0.22068</td><td colspan="2">13.057</td><td>< 2e-16 ***</td></tr>
<tr><td>Status [cognate]</td><td>-0.02754</td><td>0.17950</td><td colspan="2">-0.153</td><td>0.87808</td></tr>
<tr><td>PA (%)</td><td>0.14787</td><td>0.11885</td><td colspan="2">1.244</td><td>0.21345</td></tr>
<tr><td>Age</td><td>0.31262</td><td>0.11740</td><td colspan="2">2.663</td><td>0.00775 **</td></tr>
<tr><td>Status [cognate]:PA</td><td>-0.03704</td><td>0.07199</td><td colspan="2">-0.515</td><td>0.60690</td></tr>
<tr><th colspan="6">Random effects</th></tr>
<tr><td></td><td colspan="2">Variance</td><td colspan="3">SD</td></tr>
<tr><td>Stimulus</td><td colspan="2">2.619</td><td colspan="3">1.6185</td></tr>
<tr><td>Participant</td><td colspan="2">0.238</td><td colspan="3">0.4879</td></tr>
<tr><td>N observations</td><td colspan="5">2319</td></tr>
<tr><td>N stimulus</td><td colspan="5">120</td></tr>
<tr><td>N participants</td><td colspan="5">32</td></tr>
<tr><td>Significance codes</td><td>0 = ***</td><td>0.001 = **</td><td>0.01 = *</td><td>0.05 = .</td><td>0.1 =</td></tr>
</table>

Source: The Author (2024)

There was only one significant fixed effect in Model 4: age (corrected p-value = 0.0387). This indicates that the accuracy to decide whether a sequence of phonemes was a real word in BP was modulated by their age (β = 0.31262). Participants who were older also had more correct answers in the SLD task. However, neither cognate status of words or PA in BP had any impact in accuracy rates in this task. There was also no interaction between cognate status and PA. Thus, Hypothesis 2 and 4 were not confirmed.

In sum, Hypotheses 1 and 2 concerning cognate words were not confirmed. Cognate words across HC and BP did not have an impact on RTs or accuracy rates in either the visual or the auditory LD tasks. In addition, PA in BP only had an effect on accuracy rates in the WLD task. Hypotheses 3 is not confirmed, but Hypothesis 4 was partially confirmed. These results indicate that, for this population sample of children and teenagers speakers of HC who migrated to Brazil, cognate words do not seem to influence lexical access of isolated spoken or written words in BP. However, phonological awareness in BP does seem to positively impact the accuracy of lexical access of isolated written words in BP.

Descriptive statistics after data trimming are the following. Table 14 shows mean RTs and SDs by cognate status and task modality after data trimming. Table 15 shows mean accuracy rates and SDs by cognate status and task modality. Figures 22 and 23 show box and whiskers plots with mean RTs by cognate status and task modality.

Table 14 – Mean RTs and SDs by cognate status and task modality in Study 1

Task modality	Cognate status	Mean RTs	SDs	Ranges
SLD	Cognate	631.024	331.815	302-2492
	Noncognate	633.901	351.800	301-2441
WLD	Cognate	1029.605	414.371	318-2483
	Noncognate	1071.495	445.356	315-2499

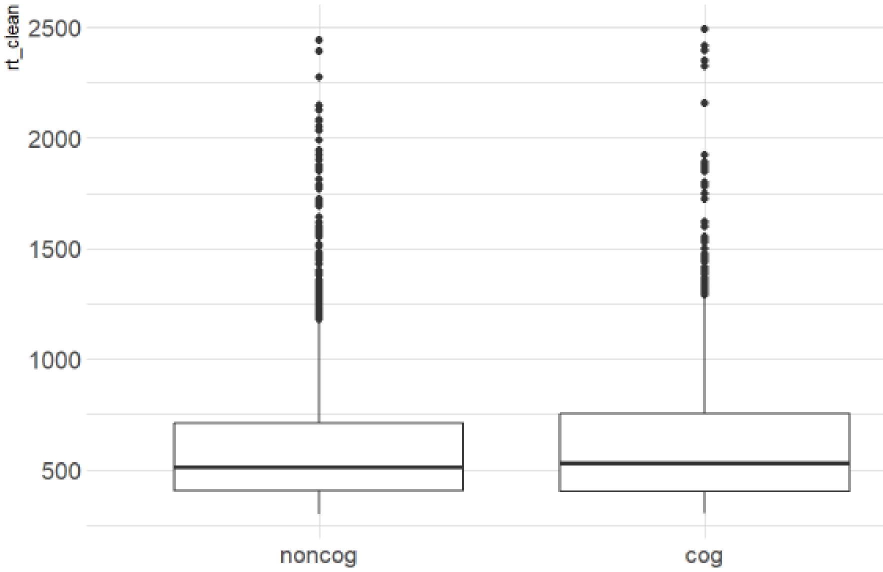
Source: The Author (2024)

Table 15 – Mean accuracy rates and SDs by cognate status and task modality in Study 1

Task modality	Cognate status	Mean accuracy rates (%)	SDs	Mean accuracy scores	Ranges
SLD	Cognate	0.869	0.338	31.844	15-47
	Noncognate	0.870	0.336	31.156	13-50
WLD	Cognate	0.925	0.263	49.880	6-59
	Noncognate	0.919	0.274	48.760	4-59

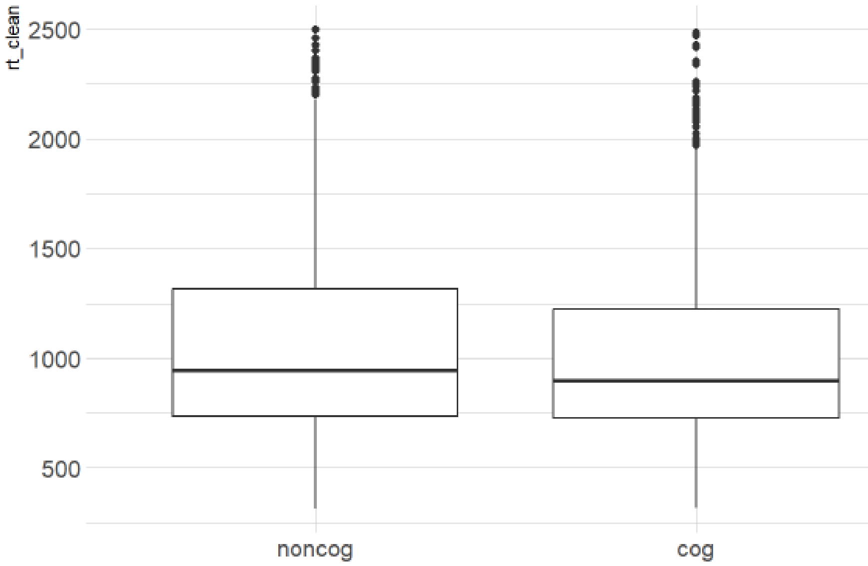
Source: The Author (2024)

Figure 22 – Mean RTs by cognate status in the SLD task in Study 1



Source: The Author (2024)

Figure 23 – Mean RTs by cognate status in the WLD task in Study 1



Source: The Author (2024)

Now I will present exploratory analyses. These include further analyses to make sure the absence of a cognate effect in planned analysis does not stem from other confounding variables. Exploratory analyses also include tests ran on residual RTs after trimming the auditory stimuli duration from RTs in the SLD task (van Dijk; Dijkstra; Unsworth, 2022). This will be detailed later.

Since age had a significant effect on RTs in the WLD task in model 1 and on accuracy rates in the SLD task in model 4, I created new subsets based on age to further explore this effect. The subsets were based on two age groups: children below 11 years old and teenagers at or above 11 years old. Teenage years are considered to start at 11 years old (Araújo *et al.*, 2022) or at 12 years old (Erlam; Philp; Feick, 2021; Nippold, 2000). In Brazil, elementary school spans from 1st to 5th grade and middle school from 6th to 9th grade; also, the number and structure of disciplines changes drastically from 5th to 6th grade. Taking this into consideration, the children age group was defined by not having reached 11 years of age yet ($N = 17$, $M_{age} = 9.06$; $SD = 0.80$) while the teenagers age group was defined by having already reached 11 years of age ($N = 31$, $M_{age} = 14.48$; $SD = 1.81$).

Four models were used to investigate the main effect of age. Model 5 had cognate status and PA as fixed effects and examined RTs for children in the WLD task. Model 6 had cognate status and PA as fixed effects and examined RTs for teenagers in the WLD task. Model 7 had cognate status and PA as fixed effects and examined accuracy rates for children in the SLD task. Model 8 had cognate status and PA as fixed effects and examined accuracy rates for teenagers in the SLD task. Tables 16 to 19 present descriptive statistics for the age groups analyses.

Table 16 – Descriptive statistics on PA and mean accuracy rates for children in the SLD task in Study 1

Cognate status	PA	Mean accuracy rates	SDs
noncog	35.897	0.846	0.366
cog	35.897	0.778	0.422
noncog	48.718	0.757	0.432
cog	48.718	0.774	0.422
noncog	61.538	0.857	0.354
cog	61.538	0.886	0.321
noncog	71.795	0.831	0.378
cog	71.795	0.833	0.376
noncog	76.923	0.875	0.334
cog	76.923	0.784	0.415

noncog	92.308	0.929	0.267
cog	92.308	0.762	0.436

Source: The Author (2024)

Table 17 – Descriptive statistics on PA and mean accuracy rates for teenagers in the SLD task in Study 1

Cognate status	PA	Mean accuracy rates	SDs
noncog	28.205	0.783	0.417
cog	28.205	0.843	0.367
noncog	38.462	0.868	0.343
cog	38.462	0.837	0.374
noncog	43.590	0.940	0.240
cog	43.590	0.882	0.325
noncog	58.974	0.769	0.427
cog	58.974	0.838	0.374
noncog	61.538	0.913	0.284
cog	61.538	0.935	0.248
noncog	64.103	0.933	0.252
cog	64.103	0.930	0.258
noncog	66.667	0.862	0.351
cog	66.667	0.897	0.310
noncog	69.231	0.906	0.296
cog	69.231	0.914	0.284
noncog	71.795	0.929	0.267
cog	71.795	0.938	0.250
noncog	76.923	0.914	0.283
cog	76.923	0.892	0.312
noncog	79.487	0.882	0.332
cog	79.487	0.889	0.320
noncog	82.051	0.815	0.396
cog	82.051	0.903	0.301
noncog	84.615	0.849	0.360
cog	84.615	0.904	0.295
noncog	89.744	0.909	0.289
cog	89.744	0.899	0.304
noncog	92.308	0.907	0.292
cog	92.308	0.920	0.272
noncog	94.872	0.905	0.296
cog	94.872	0.815	0.391
noncog	97.436	0.833	0.383

Source: The Author (2024)

Table 18 – Descriptive statistics on PA and mean RTs for children in the WLD task in Study 1

Cognate status	PA	Mean RTs	SDs
noncog	48.718	2098.500	178.521

cog	48.718	2182.333	208.406
noncog	76.923	1598.897	462.513
cog	76.923	1539.625	367.361
noncog	92.308	1470.043	406.920
cog	92.308	1503.140	401.739

Source: The Author (2024)

Table 19 – Descriptive statistics on PA and mean RTs for teenagers in the WLD task in Study 1

Cognate status	PA	Mean RTs	SDs
noncog	35.897	695.540	173.066
cog	35.897	645.420	146.159
noncog	43.590	1204.825	409.801
cog	43.590	1225.116	374.972
noncog	58.974	1036.946	454.969
cog	58.974	1030.577	409.756
noncog	61.538	901.875	257.384
cog	61.538	931.500	325.796
noncog	64.103	1047.720	476.624
cog	64.103	940.154	259.201
noncog	69.231	932.546	426.138
cog	69.231	836.509	330.269
noncog	71.795	871.842	258.586
cog	71.795	821.764	221.679
noncog	76.923	1200.283	438.884
cog	76.923	1058.041	312.232
noncog	79.487	1599.297	516.437
cog	79.487	1501.068	478.916
noncog	82.051	1190.604	461.687
cog	82.051	1043.585	393.867
noncog	84.615	1125.849	414.690
cog	84.615	1063.513	389.126
noncog	89.744	853.529	254.192
cog	89.744	844.712	234.649
noncog	92.308	1123.974	430.986
cog	92.308	1060.390	399.032
noncog	94.872	1003.608	379.982
cog	94.872	936.874	342.705
noncog	97.436	804.241	230.463
cog	97.436	836.036	364.728

Source: The Author (2024)

Tables 20 to 23 show the outputs for each model and indicated that there was no significant effect of cognate status or PA when considering age groups. In model 5, PA has a p-value equal to 0.08954.

Table 20 – Model 5 of exploratory analyses in Study 1

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.379374	0.050075	1.200201	147.365	0.00167 **
Status [cognate]	0.003287	0.027570	3.125410	0.119	0.91236
PA (%)	-0.168226	0.048634	1.750494	-3.459	0.08954
Status [cognate]:PA	-0.002149	0.028733	5.367830	-0.075	0.94309
Random effects					
	Variance		SD		
Stimulus	0.0206103		0.14356		
Participant	0.0051852		0.07201		
Residual	0.0512964		0.22649		
N observations	186				
N stimulus	104				
N participants	3				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Table 21 – Model 6 of exploratory analyses in Study 1

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.880037	0.044750	25.787683	153.744	<2e-16 ***
Status [cognate]	-0.026992	0.016835	111.687210	-1.603	0.112
PA (%)	0.022608	0.040700	19.958131	0.555	0.585
Status [cognate]:PA	-0.004359	0.005978	118.267119	-0.729	0.467
Random effects					

	Variance		SD		
Stimulus	0.0293035		0.171183		
Participant	0.0378176		0.194467		
Residual	0.0770893		0.277650		
N observations	2280				
N stimulus	120				
N participants	22				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Table 22 – Model 7 of exploratory analyses in Study 1

Fixed effects					
	Estimates	SE	z value		p
(Intercept)	2.5209	0.4107	6.138		8.34e-10 ***
Status [cognate]	-0.1740	0.2575	-0.676		0.499
PA (%)	0.2744	0.2838	0.967		0.334
Status [cognate]:PA	-0.1452	0.1722	-0.843		0.399
Random effects					
	Variance		SD		
Stimulus	3.623		1.9034		
Participant	0.279		0.5282		
N observations	544				
N stimulus	119				
N participants	7				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Table 23 – Model 8 of exploratory analyses in Study 1

Fixed effects					
	Estimates	SE	z value		p
(Intercept)	3.088665	0.248037	12.452		<2e-16 ***
Status [cognate]	0.007406	0.192508	0.038		0.969
PA (%)	0.108107	0.124568	0.868		0.385
Status [cognate]:PA	0.005124	0.083846	0.061		0.951
Random effects					
	Variance		SD		
Stimulus	2.6892		1.6399		
Participant	0.1785		0.4225		
N observations	1775				
N stimulus	120				
N participants	25				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

When selecting stimuli for the LD tasks, it was not possible to control for word length. I performed further analyses to examine whether length could have interacted with cognate status and lead to the absence of cognate effect. Model 9 had cognate status and word length as fixed effects and tested their effects on RTs in the WLD task. Model 10 tested the same fixed effects on RTs in the SLD task. Model 9 showed a significant effect of length in RTs in the WLD task ($p < 1.5^{e-13}$), but no interaction with cognate status (Table 24). Word length is known to influence word processing: longer words take longer to process than shorter words (Cortese; Balota, 2012; Yap; Balota, 2015). This effect may be more intense for readers who have less efficient decoding skills (Gerth; Festman, 2021; Barton *et al.*, 2014). Word length could have obscured the cognate effect if only shorter words had shown differences in RTs between cognate and noncognate ones.

Table 24 – Model 9 of exploratory analyses in Study 1

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.555331	0.073563	60.425833	89.112	<2e-16 ***
Status [cognate]	0.058734	0.048162	104.419083	1.220	0.2254
Length	0.059138	0.006982	106.276924	8.470	1.5e-13 ***
Status [cognate]:Length	-0.012670	0.006978	106.013607	-1.816	0.0722
Random effects					
	Variance		SD		
Stimulus	0.0158010		0.125702		
Participant	0.0772285		0.277900		
Residual	0.0752769		0.274366		
N observations	2466				
N stimulus	120				
N participants	25				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 10 showed no significant effect of length and no interaction between cognate status and length (Table 25). There was a marginal main effect of length ($p = 0.052$) which was discarded because the p-value would increase when adjusted. Thus, it seems improbable that word length may have obscured the cognate effect. Tables 26 and 27 present mean RTs for each word length in both task modalities.

Table 25 – Model 10 of exploratory analyses in Study 1

Fixed effects

	Estimates	SE	df	t	p value
(Intercept)	6.418794	0.055864	123.746148	114.901	<2e-16 ***
Status [cognate]	0.058253	0.049552	103.762147	1.176	0.2424
Length	-0.014961	0.007701	113.081566	-1.943	0.0545
Status [cognate]:Length	-0.008717	0.007692	112.660533	-1.133	0.2595
Random effects					
	Variance		SD		
Stimulus	1.102e-02		0.104983		
Participant	2.107e-02		0.145140		
Residual	1.643e-01		0.405325		
N observations	2016				
N stimulus	120				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Table 26 – Mean RTs and SDs for each word length in the SLD in Study 1

Word length in the SLD	Mean RTs	SDs	Ranges
3	653.831	373.215	310-2158
4	619.555	325.053	301-2395
5	654.334	346.622	301-2415
6	638.507	313.189	301-2125
7	653.456	400.393	301-2441
8	625.990	368.741	303-2492
9	575.356	267.111	309-1593
10	544.229	268.651	303-1875

Source: The Author (2024)

Table 27 – Mean RTs and SDs for each word length in the WLD in Study 1

Word length in the WLD	Mean RTs	SDs	Ranges
3	880.370	325.755	315-2168

4	921.840	383.186	473-2483
5	988.552	388.981	318-2458
6	991.348	409.343	436-2427
7	1052.142	428.671	406-2479
8	1114.873	456.728	441-2499
9	1200.533	458.414	459-2499
10	1413.317	447.488	633-2426
11	1203.842	378.232	582-2403
12	1336.314	487.214	595-2334

Source: The Author (2024)

Descriptive statistics in Tables 26 and 27 show that RTs varied as function of word length much more in the WLD task than in the SLD task. This difference is expected due to the processing pathways involved in oral and written word recognition (Harm; Seidenberg, 2004; Li; Zhao, 2013; Li; Zhao; MacWhinney, 2007; Seidenberg; McClelland, 1989). In addition, the longest RTs in the WLD task are associated with longer words, especially 10-letter words. There were two 11-letter words and two 12-letter words in the stimuli for the WLD task which did not have length-wise counterparts in the SLD task, but their mean RTs were not longer than for the 10-letter words. Despite the main effect of length in model 9, there was no interaction with cognate status in any of the models 7-10. So cognate status did not influence RTs in relation to words of different lengths, and word length did not influence the cognate effect.

Moreover, another possible reason for the absence of the cognate effect could be weak word decoding and recognition skills. The Letter Identification scores collected in this study may be associated to participants' word recognition abilities. Indeed, it has been shown that letter knowledge predicts word recognition (Catts *et al.*, 2015; Hogan; Catts; Little, 2005; Muter *et al.*, 2004; Oliveira; Germano; Capellini, 2016). This way I ran models 11 and 12 of accuracy rates with cognate status and letter identification as fixed effects. Tables 28 and 29 present the output for each model.

Table 28 – Model 11 of exploratory analyses in Study 1

Fixed effects				
	Estimates	SE	z value	p
(Intercept)	4.23531	0.37979	11.152	<2e-16 ***
Status [cognate]	0.24250	0.25469	0.952	0.341
LI (%)	-0.12461	0.23018	-0.541	0.588

PA (%)	0.52215	0.24770	2.108		0.035 *
LI:PA	-0.07486	0.18046	-0.415		0.678
Status [cognate]:LI	0.01763	0.08275	0.213		0.831
Random effects					
	Variance		SD		
Stimulus	5.044		2.2458		
Participant	0.896		0.9466		
N observations	2675				
N stimulus	120				
N participants	25				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Table 29 – Model 12 of exploratory analyses in Study 1

Fixed effects				
	Estimates	SE	z value	p
(Intercept)	2.82825	0.23427	12.073	<2e-16 ***
Status [cognate]	-0.01425	0.17951	-0.079	0.9367
LI (%)	0.08781	0.14886	0.590	0.5553
PA (%)	0.25151	0.15106	1.665	0.0959
LI:PA	0.13043	0.13549	0.963	0.3357
Status [cognate]:LI	0.06446	0.07291	0.884	0.3766
Random effects				
	Variance		SD	
Stimulus	2.6264		1.6206	

Participant	0.3127		0.5592		
N observations	2319				
N stimulus	120				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

However, there was no main effect nor interaction involving the letter identification scores in either task modality. Letter identification was not associated either with cognate status nor with PA in BP. According to model 11 (Table 28), it might be that PA has more impact on word recognition for this sample of participants. Also, it does not seem that the cognate effect was modulated by the letter and grapheme-phoneme conversion knowledge tapped by the Letter Identification test.

Now I will describe further exploratory analyses which were inspired by the ones performed by van Dijk, Dijkstra, and Unsworth (2022). They used a self-paced sentence listening task to investigate syntactic cross-language influences. Since the segments participants listened to had different durations, they followed the procedures recommended by Marinis (2010). I performed further exploratory analyses using this procedure on the SLD task because word length of stimuli varied.

First, residual RTs were computed. For this, the duration of each audio file was removed from its respective RT for all participants so that the residual RT would indicate the time participants took to respond after the end of the audio. This means that if a participant answered before the end of the audio, the residual RT would be a negative number. In order to solve this, a constant was added to all RT values. This constant was the lowest RT value transformed to a positive number plus one. The lowest RT value for each participant was calculated; it was -1,221 ms. This number was transformed into a positive number (1,221) and was added one (1,222). This constant was aggregated to all residual RTs in both task modalities.

After calculating the residual RTs in the SLD and adding the constant to both tasks — so that proportions for data visualization would be the same —, data trimming was carried out similarly to the steps reported in the planned analyses above. The only descriptive statistic which changed from planned analyses to this residual RTs one is RT values. Table 30 presents mean residual RTs, SDs, and ranges for both lexical decision tasks.

Table 30 – Mean residual RTs, SDs, and ranges for both lexical decision tasks in Study 1

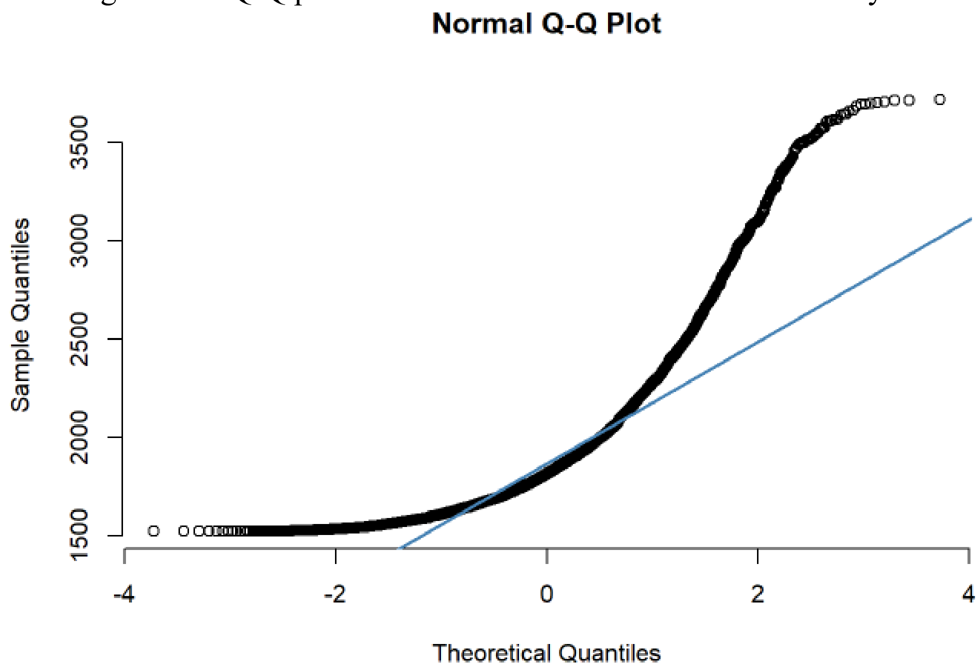
Task	Mean response times (ms)	SDs	Range (ms)
Spoken Lexical Decision task	761.416	2871.546	1—219557
Written Lexical Decision task	2449.968	11516.810	1223—1174253

Source: The Author (2024)

Data from participants who had accuracy rates lower than 70% was removed. RT values deviating 2.5 SDs below or above the mean for each participant and RT values shorter than 1522 ms (300 ms + 1222) and longer than 3722 ms (2500 ms + 1222) were also removed. Data was also removed according to scores in the HC vocabulary test and the Letter Identification test. These steps removed the same percentage of the data as in the planned analyses.

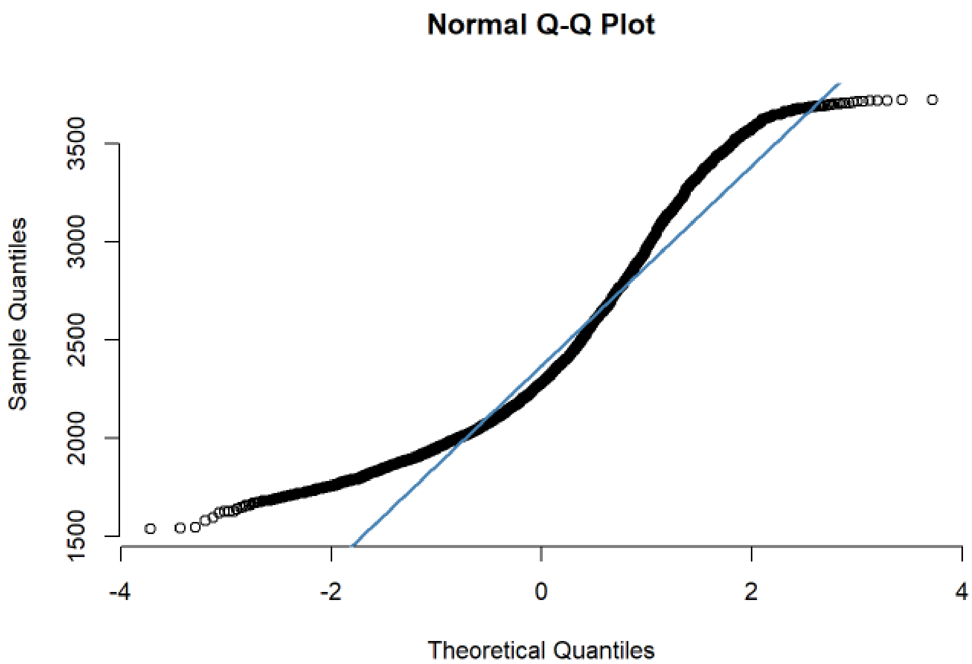
The distribution of RTs was also not normal (SLD task: $A = 247.21$, $p < 0.001$; WLD task: $A = 120.79$, $p < 0.001$), which can be verified with Figures 24 to 27 which show Q-Q plots and histogram plots.

Figure 24 – Q-Q plot for residual RTs in the SLD task in Study 1



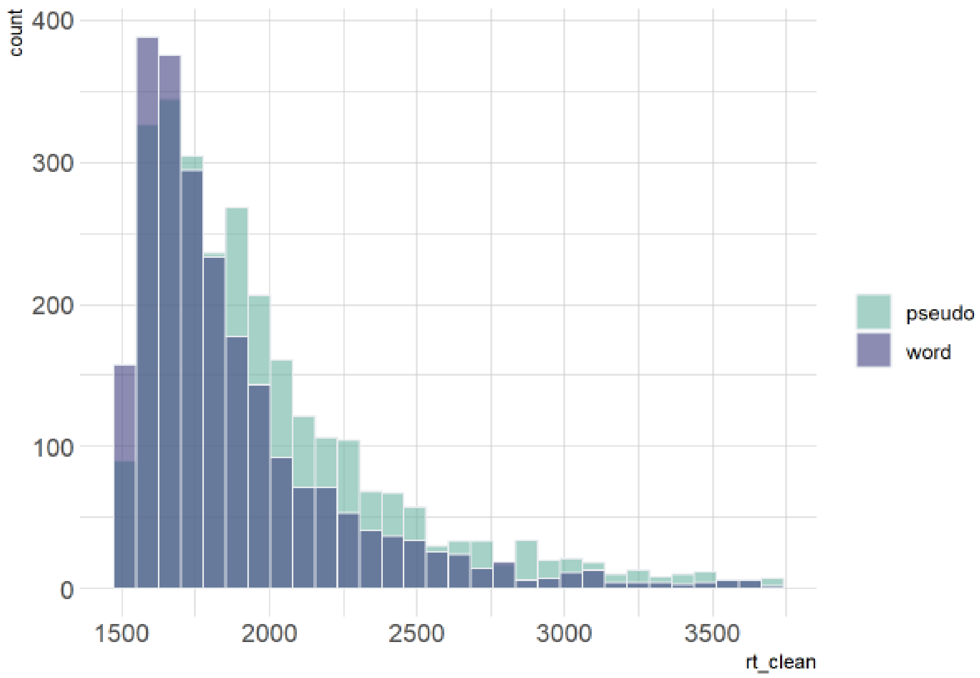
Source: The Author (2024)

Figure 25 – Q-Q plot for residual RTs in the WLD task in Study 1



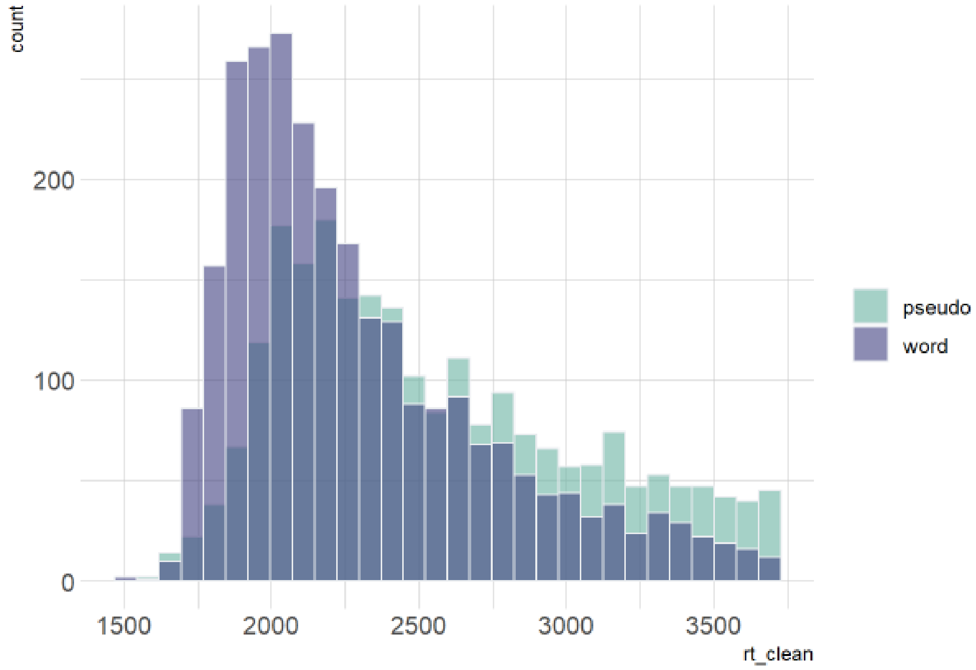
Source: The Author (2024)

Figure 26 – Histogram plot for residual RTs in the SLD task in Study 1



Source: The Author (2024)

Figure 27 – Histogram plot for residual RTs in the WLD task in Study 1



Source: The Author (2024)

Thus, I used a Wilcoxon Ranks Sum test to examine whether there was a significant difference in RTs between real word trials and pseudoword trials in both LD task modalities. Indeed, RTs to real words were shorter than to pseudowords (SLD task: $W = 3744455$, $p < 0.001$; WLD task: $W = 4042029$, $p < 0.001$). Descriptive statistics may be examined in Table 31.

Table 31 – Mean RTs and SDs for words and pseudowords in both LD tasks in Study 1

	Trial type	Mean RTs	SDs	Ranges
SLD	Word	1879.544	366.522	1523-3714
	Pseudoword	1986.133	418.025	1523-3715
WLD	Word	2297.039	446.381	1537-3721
	Pseudoword	2543.335	507.354	1546-3719

Source: The Author (2024)

Before removing data from participants who had accuracy scores lower than 70% in the LD tasks, the mean age was 12.56. After that, mean age for the SLD data was 12.60 and for the WLD data was 13.23. Median age (13) was the same before and after accuracy trimming for both tasks.

Contrasts were set and variables were centered following the steps in the planned analyses. Two mixed-effect linear models were used to test Hypotheses about RTs in relation to residual RTs. They were constructed exactly like in the planned analyses described above.

Model 1 had the following code: `m11r <- lmer(data = s1W70_wordcorrect, log(rt_clean) ~ status + PApcent_s + age_s + status*PApcent_s + (1 + status | participant) + (1 | stimulus))`. Table 32 presents estimates, standard errors, and p-values. Results show that only age had a significant main effect on RTs in the WLD task ($p = 0.00254$).

Table 32 – Model 1 on residual RTs in the WLD task in Study 1

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.738212	0.025483	25.481175	303.661	<2e-16 ***
Status [cognate]	-0.013291	0.007957	111.827931	-1.670	0.09764
PA (%)	-0.021960	0.023475	21.923265	-0.935	0.35973
Age	-0.067693	0.020056	23.686669	-3.375	0.00254 **
Status [cognate]:PA	-0.000839	0.002834	65.883337	-0.296	0.76812
Random effects					
	Variance		SD		
Stimulus	6.562e-03		0.08101		
Participant	1.447e-02		0.12028		
Residual	1.635e-02		0.12788		
N observations	2466				
N stimulus	120				
N participants	25				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 2 had the following code: `m12r <- lmer(data = s1S70_wordcorrect, log(rt_clean) ~ status + PApcent_s + age_s + status*PApcent_s + (1 + status | participant) + (1 | stimulus))`.

Table 33 presents estimates, standard errors, and p-values. Results show that there was no significant main effect or interaction in the model for RTs in the SLD task.

Table 33 – Model 2 on residual RTs in the SLD task in Study 1

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.507e+00	1.069e-02	3.499e+01	702.466	<2e-16 ***
Status [cognate]	9.815e-04	4.877e-03	1.008e+02	0.201	0.841
PA (%)	-5.642e-03	1.051e-02	2.839e+01	-0.537	0.596
Age	-1.161e-02	1.014e-02	2.919e+01	-1.145	0.261
Status [cognate]:PA	-3.365e-04	3.417e-03	3.646e+02	-0.098	0.922
Random effects					
	Variance		SD		
Stimulus	1.391e-03		0.037297		
Participant	2.861e-03		0.053485		
Residual	2.241e-02		0.149703		
N observations	2016				
N stimulus	120				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

When comparing the results from models 1 and 3 in the planned analyses and from models 1 and 2 in the residual RTs exploratory analyses, no difference in main effects or interactions is found. In sum, exploratory analyses with residual RTs did not add any different results from the ones presented in planned analyses. Table 34 show mean residual RTs and SDs by cognate status and task modality after data trimming. Finally, in this study, it seems that children and teenager speakers of HC who migrated to Brazil and use BP in their everyday lives do not demonstrate effects from cognate words across HC and BP when they are reading or listening to isolated words in BP. However, they do display to be positively influenced by higher scores on a BP PA test when recognizing words in BP.

Table 34 – Mean residual RTs and SDs by cognate status and task modality in Study 1

Task modality	Cognate status	Mean RTs	SDs	Ranges
SLD	Cognate	1855.901	351.800	1524-3714
	Noncognate	1853.024	331.815	1523-3663
WLD	Cognate	2293.495	445.356	1540-3705
	Noncognate	2251.605	414.371	1537-3721

Source: The Author (2024)

In the next section, I will discuss these results in relation to the literature and to potential alternative perspectives.

3.3 DISCUSSION

In this section, I will refer to the research questions, objectives and hypotheses and sum up the results presented in Section 3.2 before discussing them in light of the literature.

Study 1 investigated whether children and teenagers who are native speakers of HC show effects of cross-linguistic interaction between HC and BP on lexical access during comprehension of spoken and written BP. The study was guided by the following research questions:

1. Do cognate words across HC and BP, in comparison with noncognate words, facilitate visual and auditory lexical access in BP?
2. Does phonological awareness (PA) in BP influence visual and auditory lexical access in BP?

In order to answer these research questions, Study 1 pursued the following specific objectives:

1. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during visual and auditory lexical decision tasks in BP;

2. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce higher accuracy rates during visual and auditory lexical decision tasks in BP;
3. To investigate whether higher PA scores in BP, in comparison with lower PA scores, produce shorter RTs in visual and auditory lexical decision tasks in BP;
4. To investigate whether higher PA scores in BP, in comparison with lower PA scores, produce higher accuracy rates in visual and auditory lexical decision tasks in BP.

Objectives 1 and 2 are related to research question 1, while Objectives 3 and 4, to research question 2. The Hypotheses for Study 1 are:

1. Cognate words, in comparison with noncognate words, produce shorter RTs during both visual and auditory lexical decision tasks in BP.
2. Cognate words, in comparison with noncognate words, produce higher accuracy rates during both visual and auditory lexical decision tasks in BP.
3. Higher scores in PA in BP, in comparison with lower PA scores, produce shorter RTs in both visual and auditory lexical decision tasks in BP.
4. Higher scores in PA in BP, in comparison with lower PA scores, produce higher accuracy scores in both visual and auditory lexical decision tasks in BP.

In order to test these hypotheses, I invited children and teenage speakers of HC who were living in Brazil to perform 4 tasks in BP and one in HC. A total of 48 participants completed one visual and one auditory lexical decision (LD) task with cognate and noncognate words across HC and BP, one phonological awareness (PA) test in BP, one letter and grapheme-phoneme conversion identification (LI) task, and one HC receptive vocabulary test. The lexical decision tasks were intended to test whether cognate words across HC and BP would have any significant influences in the lexical access of words in BP recognized visually and auditorily. The PA test was intended to measure phonological awareness to verify whether lexical access would be modulated by it.

Results from planned analyses showed no significant main effects or interactions involving cognate status. There were marginal effects which were not considered because of Bonferroni adjustments to p-values. This indicates that the presence of cognate words across HC and BP did not facilitate and did not inhibit lexical access of words in BP. On the other hand, there was a significant and positive main effect of PA in BP on accuracy rates during the

WLD task even after adjusting p-values ($p = 0.04$). This indicates that participants who scored higher in the PA test also had higher accuracy scores in the WLD task.

Now I will summarize results from the exploratory analyses. Planned analyses revealed a significant effect of age on RTs in the WLD task ($p = 0.0225$, adjusted) and on accuracy rates in the SLD task ($p = 0.0387$, adjusted). Descriptive statistics showed that older participants presented shorter RTs and higher accuracy scores than younger ones (Tables 16 to 19). These differences may be expected considering the linguistic development of participants: the older they are, the better reading skills and vocabulary will be. However, since the sample of participants in this study ranged from 7 to 17 years of age, I further examined the main effects of age with four exploratory models: one on RTs in the WLD for children, one on RTs in the WLD for teenagers, one on accuracy rates in the SLD task for children, and one on accuracy rates in the SLD for teenagers. These models had cognate status and PA as fixed effects. There was no significant main effect or interaction in any of these models. This may indicate that the absence of a cognate effect is not caused by the age range of participants.

Exploratory analyses included an investigation about the effects of word length and about a potential effect of the Letter Identification (LI) task scores. The exploratory models on word length will be presented in Section 3.3.2. Meanwhile, I will present the exploratory analysis on letter identification.

As mentioned in Section 3.2.2, letter identification/knowledge has been shown to predict word recognition (Catts *et al.*, 2015; Hogan; Catts; Little, 2005; Muter *et al.*, 2004; Oliveira; Germano; Capellini, 2016). Then, two exploratory models were run to consider influences of an objective measure of word recognition skill on accuracy rates in both LD tasks. However, there was no significant main effect or interaction of the scores in the LI task in either model. This might indicate that knowing letters and grapheme-phoneme conversion had little influence in the LD tasks or that the LI task did not measure the construct it was intended to test. Considering that data inclusion in the inferential analyses also used a minimum score threshold in the LI task, it is possible that letter knowledge did not obscure any cognate effects in the LD tasks.

Exploratory analyses included a data trimming strategy used by van Dijk, Dijkstra, and Unsworth (2022). The goal of this strategy was to calculate residual RTs for the SLD task, that is, the time participants took to respond to the auditory stimuli subtracting the duration of each auditory stimuli. After conducting this extra trimming, the same models used in planned analyses were run. However, these exploratory analyses presented no significant differences in relation to the planned ones.

Thus, results showed that, in this study, it seems that children and teenager speakers of HC who migrated to Brazil and use BP in their everyday lives do not demonstrate effects from cognate words across HC and BP when they are reading or listening to isolated words in BP. However, they do display to be positively influenced by higher scores on a BP PA test when recognizing words in BP. This means that Hypotheses 1, 2 and 3 were not confirmed. Hypothesis 4 was confirmed only in relation to the visual lexical decision task, not the auditory one.

I will first discuss the disconfirmation of Hypothesis 3 and the confirmation of Hypothesis 4 in relation to the WLD task and later discuss the disconfirmation of Hypotheses 1 and 2 and the absence of cognate effects. I will review the two last hypotheses first although it is not the most logical order because these last hypotheses are related to the only significant effect seen in Study 1.

Hypothesis 3 stated that higher scores in PA in BP are associated with shorter RTs in both visual and auditory lexical decision in BP. This was expected because PA is correlated with word decoding (Swank; Catts, 1994) and predicts word reading years later (Hogan; Catts; Little, 2005). Thus, higher PA scores would be associated with faster lexical access and shorter RTs in the LD tasks. However, PA displayed no significant effect on RTs in either task modality. This result is in contrast with studies pointing to the association between PA and word processing, although most studies present evidence with accuracy rates. In the present study, response times varied vastly as SDs on Tables 14 and 15 show. It is possible that PA effects on RTs might have not emerged due to this large variance in RTs. On the other hand, PA did present a significant effect on accuracy rates in the WLD task.

Hypothesis 4 stated that higher scores in PA in BP are associated with higher accuracy scores in both visual and auditory lexical decision in BP. This was true for the written LD only ($\beta = 0.517175$, $p < 0.001$). This result is in line with a vast body of research showing that there is a mutual relationship among PA and language and reading development, even across different languages (Amorim *et al.*, 2020; Caravolas; Bruck, 1993; Muter *et al.*, 1998; Paula; Mota; Keske-Soares, 2005; Toffoli; Lamprecht, 2008; Rezaei; Mousanezhad Jeddi, 2020; Santos; Befi-Lopes 2012; Verhoeven, 2000; Verhoeven; van Leeuwe, 2012; Verhoeven; Voeten; Vermeer, 2019; see more in Defior, 2004).

For instance, Muter *et al.* (1998) conducted a longitudinal experimental study which showed that letter knowledge and segmentation influenced spelling and reading development in the first year of schooling. Spelling was predicted by phonological awareness throughout the 2 years and by rhyming in the last year. Paula, Mota, and Keske-Soares (2005) observed that an intervention with PA and grapheme-phoneme conversion training contributed to improving the reading and writing skills of 76,47% of children in the experimental group in relation to the control group. Santos and Befi-Lopes (2012) observed that PA correlated moderately and negatively with spelling errors for 9- and 10-year-olds. Rezaei and Mousanezhad Jeddi (2020) observed a direct and negative effect of phonological awareness on reading errors for 209 children. In addition, Amorim, Jeon, Abel, Felisberto, Barbosa, and Dias (2020) performed an experimental study on the effectiveness of an educational game for 749 children. The game trained them on phonological awareness, word reading, and writing skills. The experimental group improved 68% in reading scores and 48% in writing scores in comparison with the control group after using the game for 3 months.

In the present study, there was no influences of PA scores on the spoken LD task. Oral language skills are shown to be associated with PA (Caravolas; Bruck, 1993; Cooper *et al.*, 2002; Hipfner-Boucher *et al.*, 2014) — for a review, see Anthony and Francis (2005). But for the population sample of the present study, the relationship between BP listening skills and BP PA might not be mutual. It is relevant to refer back to Saiegh-Haddad (2019) and her definition of L2 PA. She proposes that L2 PA is formed of a metalinguistic component and of a linguistic one: the first one equates to other definitions of PA while the second one is similar to the concept of lexical quality. I would like to highlight that Saeigh-Haddad (2019) proposes that L2 PA is also composed of L2 high-quality lexical representations, because this aspect will be important for a speculative explanation of the absence of cognate effects in this study. This could indicate that, for the sample of HC-speaking children and teenagers of the present study, listening and reading skills may be contributing to develop their PA skills, but PA has more impact on the mediation between orthography and meaning via the phonological code during reading.

Thus, Hypothesis 4 was partially confirmed, which indicates that, for this sample of HC-speaking children and teenagers, PA in BP is associated with their accuracy in deciding whether strings of letters are real words in BP but not whether sequences of phonemes are real words in BP. Next, I will discuss Hypotheses 1 and 2, which were not confirmed.

Hypotheses 1 and 2 stated that cognate words, in comparison with noncognate words, are associated with shorter RTs and with higher accuracy rates during both visual and auditory lexical decision in BP. The absence of any cognate effect — either a facilitatory or an inhibitory one, either in RTs or in accuracy rates — diverges from most of the evidence reported in the literature. Although the majority of studies investigating cross-language activation effects via cognate words focuses on adult participants, most studies with children also indicate cognate facilitation effects.

I performed two brief searches in the literature available at the main database at Periódicos da CAPES. In the first one, I used the following keywords: cognate children language activation. After that, I examined the abstracts of the 16 studies retrieved and selected the ones which reported quasi-experimental linguistic investigations. I have found 9 studies looking into cross-language activation via cognate words with children (Arêas da Luz Fontes *et al*., 2021; Bosma *et al.*, 2019; Bosma; Nota, 2020; Brenders; van Hell; Dijkstra, 2011; Davis; Bowman; Kaushanskaya, 2018; Gastmann; Poarch, 2022; Jared *et al.*, 2012; Schröter; Schroeder, 2016; Woolpert, 2019). Seven of them showed cognate facilitation effect (Arêas da Luz Fontes *et al*., 2021; Bosma *et al.*, 2019; Bosma; Nota, 2020; Brenders; van Hell; Dijkstra, 2011; Gastmann; Poarch, 2022; Jared; Cormier; Levy; Wade-Woolley, 2012; Schröter; Schroeder, 2016) while only two showed inhibition or no effect at all (Davis; Bowman; Kaushanskaya, 2018; Woolpert, 2019).

In the second search, I used the following keywords: cognate language activation minority. I have found 7 studies investigating cross-language activation via cognate words whose participants were speakers of minority languages (Bosma *et al*., 2019; Bosma; Nota, 2020; Campos, 2023; Davis; Bowman; Kaushanskaya, 2018; Kirk *et al.*, 2018; Muntendam *et al*., 2022; Woolpert, 2019). Among these 7 studies on minority languages, 5 studies focused on children and/or teenagers (Bosma *et al.*, 2019; Bosma; Nota, 2020; Campos, 2023; Davis; Bowman; Kaushanskaya, 2018; Woolpert, 2019). Among these 5, two studies reported cognate facilitation (Bosma *et al.*, 2019; Bosma; Nota, 2020), one reported inhibition (Davis; Bowman; Kaushanskaya, 2018), one reported no effect (Woolpert, 2018), and one could not be accessed (Campos, 2023). Therefore, in line with a review by Squires *et al*. (2020), most studies with children show evidences of cognate words leading to facilitated language processing.

Hypotheses 1 and 2 in this study expected to observe a cognate facilitation effect in both lexical decisions tasks just like in Brenders, van Hell, and Dijkstra (2011), Jared *et al.* (2012), Schröter and Schroeder (2016), Bosma *et al.* (2019), Arêas da Luz Fontes *et al.* (2021), Bosma and Nota (2020), and Gastmann and Poarch (2022). Hypotheses 1 and 2 implied that lexical access for bilingual or multilingual people is non-selective. This was based on a vast literature on the bilingual lexicon and on predictions from bilingual language processing models. For example, BIA+ states that "[t]he information flow in bilingual lexical processing proceeds exclusively from the word identification system toward a task/decision system on the activation state of words" (Dijkstra, 2005, p. 197). This means that any top-down influences have no impact on the activation of words and their levels of representation. In addition, "cognate effects persist even when strong language membership cues are present in the absence of overlap between orthographies" (Winther; Matusevych; Pickering, 2023, p. 113). So even in a monolingual context, the bilingual person would still be affected by language co-activation.

Also, according to the BIA+ model, cognate words would be activated in parallel independently of whether the bilingual person was aware of form and meaning similarities between these words across languages. Davis, Bowman, and Kaushanskaya (2018, p. S23) mentioned that "explicit training on how to utilise cognates in the service of reading comprehension" could be useful. Indeed, Hipfner-Boucher *et al.* (2016) showed that, since grade 1, children may demonstrate being aware of cognates. However, intervention for raising awareness of cognate words does not seem to help learning these words (Otwinowska *et al.*, 2020). Thus, the lexicons from both languages seem to be accessed at the same time via cognate words in a non-selective manner, and the awareness of cognate words does not appear to influence their processing.

Contrarily, results from this study are not in line with the majority of the existent body of evidence indicating that cognate words are accessed in a non-selective way. The results of the present study do not point to any facilitation effect and do not suggest any reliable interference effect from cognate words. Thus, these null results could be in accordance with a selective view of lexical access. Dijkstra (2005) explains null results in relation to two possibilities of phenomena. The first possibility would be that "the item types were not really comparable or were not matched properly" (Dijkstra, 2005, p. 182). The second one would be either that there was not enough relative activation of one of the languages or that participants would not always respond to according to task instructions (Dijkstra, 2005, p. 182). I will next discuss the results considering these possibilities.

3.3.1 Discussing stimuli

There were two important limitations in terms of stimuli selection for this study. The first one was frequency of occurrence and the second one was length.

I could not find an annotated and normalized corpus of HC with annotations for parts of speech from which to extract the stimuli word. Instead, I compiled, translated, annotated and organized by estimated frequency HC words based on a HC-English dictionary and a HC-BP glossary. All these steps were guided and revised by an informant who was a highly-educated native speaker of HC who was living in Brazil since 2014 and who was enrolled in a Brazilian university course. The informant organized all 5,617 words according to his perceived frequency of occurrence. A total of 5,291 of them were categorized as being highly frequent, and this was the original set of words from which I selected cognate and noncognate words.

I understand that perception is a subjective measure and that perception of frequency does not equal actual frequency. I also understand that the frequency estimates originated from the perception of only one person. These frequency estimates would have been more reliable if they had been gathered from a bigger number of informants. I did find another comparable informant; however, they had very limited time available and could not contribute. Thus, it is clear that the frequency estimates in this study may be questioned.

The second limitation was word length. It was not possible to keep it constant due to the number of cognate words available from the final word set. Consequently, stimuli words, both real ones and pseudowords, varied in length from 3 to 12 characters, and participants' RTs varied according to word length. Exploratory analyses (Tables 24 and 25) showed that longer words presented longer RTs in the WLD task ($p < 0.001$). As mentioned in Section 3.2, there was no main effect of length for RTs in the SLD task and no interaction between cognate status and word length, especially after correcting p-values (*p*s > 0.1). That is, the fact that longer words were also cognate did not facilitate (nor inhibit) their processing in comparison with noncognate ones. For this reason, I believe the fact that stimuli varied in word length did not obscure the cognate effect for this sample of participants.

It was also possible to observe that RTs for the SLD task were shorter than for the WLD. This difference of RTs in task modality processing is not surprising and is predicted by language processing models. Children and teenagers would probably be exposed to BP first in the spoken modality and later in the written one. Recognizing spoken words would take fewer processing steps than recognizing written words when decoding is still not automatized: while the input for the spoken word is the phonological information, the input for the written one is the orthographic information. The DevLexII model proposes that activating the phonological information of a word leads to its semantic information (see Figure 6). So, in speech processing in monolinguals or bilinguals, the majority of activation would be directed first from phonology to meaning. Meanwhile, during written word recognition, both the BIA+ model and the Triangle Model allow that the activation flows from orthography to phonology and then to semantics and from orthography to semantics directly (see Figures 5 and 3 respectively). This does not mean the activation spread is restricted to one pathway; it may spread both ways, but one of them ends up having more effective results in fulfilling the task. It is possible that children and teenagers in this study were relying mostly on the indirect route via phonology to access words. This would lead to the (at least numerical) differences in RTs across the SLD task and the WLD task seen in Table 14. In addition, children displayed longer RTs overall than teenagers in the WLD task (Tables 18 and 19), which could indicate that children were still automatizing their decoding skills in BP in relation to teenagers.

Another limitation in relation to stimuli was the fact that only in the WLD there were 3 identical cognates: *liberal*, *matinal*, and *mal*. Identical cognates usually show the most intense cognate effect. However, although there were 3 identical cognates in the stimuli list for this task, no reliable cognate effect was seen.

3.3.2 Discussing language activation

In this section, I will consider the explanation for null results posited by (Dijkstra, 2005, p. 182) about the relative activation of one of the languages of participants. There was no indication of language co-activation either in the SLD or the WLD tasks. It would be possible to argue that the lack of a cognate effect in the written task could stem from weak word decoding skills and vocabulary. However, if poor reading skill were the reason for null results, cognate effects could have emerged in the spoken task. In this case, lack of vocabulary or linguistic knowledge could explain the absence of cognate effects. This might indicate that the parallel activation of languages was not visible in either RTs or accuracy rates due to more general language processing aspects.

Accuracy rates in the tasks varied vastly (Table 15). This meant that more than 50% of the data was removed from both LD tasks for inferential analyses. It is important to add that ambient noise was a constant limitation in this study. Most data collection sessions occurred in school rooms which were quiet enough for the participants to perform the tasks using headphones or earplugs. However, there was always noises from children playing outside and people talking, which were unavoidable. This might have distracted some participants. Although there was no task or section in a task to test participants' attentional level during data collection, it was not hard to tell when participants were getting distracted from task because they were always accompanied by the researcher during data collection. The protocol was to take notes when participants responded too quickly and when participants looked away from the screen. In cases when this happened many times during a task, data would not be added to the inferential analyses. Due to these limitations and compensations, the variation in accuracy is probably originated in participants' linguistic knowledge and language skills.

3.3.3 Discussing information about participants

Now I will consider the information about this population sample collected via questionnaire. In this study, participants who answered the language history questionnaire in full (*N* = 40) reported having good proficiency in HC (5.28) and almost very good proficiency in BP (5.80) (Table 4). Their answers for hours using each language per day indicated less variation for the use of HC and more variation for the use of BP. This can be observed in the smaller SDs for hours using HC in comparison with SDs for hours using BP (Table 5).

Most participants (*N* = 33) answered the LHQ themselves with help from the researcher. During this step, they made some comments about their conversation habits. For example, some of them mentioned that they did not really talk much in school, neither with teachers nor classmates. They explained they did not have many friends to engage in conversation with. They added that since their parents or family worked all day long, they were used to being alone at home. So, they did not have many social interactions at home other than conversations via phone chat applications. These comments might indicate that these participants have few opportunities or little motivation to engage in in-person conversations either in HC e or BP. However, they did report in the LHQ to use BP more than HC in activities. HC was used more in virtual interactions according to answers to the questionnaire.

Thus, although participants reported having good levels of proficiency in BP, this self-reported proficiency was not reflected in their accuracy rates. The self-perception on language abilities may be influenced by their communicative successes in everyday language use situations. However, these levels of perceived proficiency may not have been not sufficient to reach 70% of correct answers in the tasks completed in the studies or these levels may not equal knowing the vocabulary present in the tasks. In addition, the impact of proficiency on language co-activation and on the cognate effect seems to follow an inverted U-shaped model: in order for these effects to emerge, the bilingual person must be between a minimum level of proficiency in the L2 and a maximum one (Bultena; Dijkstra; van Hell, 2014). In other words, having too little knowledge of the language or having too much knowledge/proficiency will not produce co-activation effects. Therefore, one possibility for explaining the null results in this study is that there was not enough cross-language activation due to lack of linguistic knowledge in order to cognate words to have any impact on RTs or accuracy rates.

I was able to find two studies interested in the effects of cognate words on language processing by bilingual children who spoke a heritage language. The first one was Davis, Bowman, and Kaushanskaya (2018) and the second one was Woolpert (2018).

Davis, Bowman, and Kaushanskaya (2018) investigated whether the presence of cognate words in a story would have an impact on children's reading aloud. Participants were monolingual English speakers and simultaneous Spanish-English bilinguals between 7 and 13 years of age. Bilinguals read the story which contained cognate words more slowly and made more errors in comparison with the story which did not have cognate words. Monolinguals displayed shorter latencies than bilinguals and demonstrated no differences across stories. This result for bilinguals is not in accordance with most of the literature, which shows a cognate facilitation effect even in language production (Preuss, 2012; Poarch; van Hell, 2014).

This inhibition effect may be caused by characteristics of the sample of participants. Davis, Bowman, and Kaushanskaya calculated SES of participants' families according to years of parents' education. There was a significant difference between groups in relation to SES: bilinguals belonged to lower SES families than monolinguals. This characteristic may be associated with the language development of this bilingual children group (Piccolo *et al.*, 2013).

Woolpert (2018) tested the effect of cognate words and false friends on Spanish-English 2nd- to 4th-grade children reading skills. Most of these children were immigrants from Mexico, so English was their majority language while Spanish was the minority one. They completed decoding, vocabulary and reading comprehension tasks in both languages. Cognate words had no impact on any measures, and false friends caused interference both in accuracy and in RT measures.

Woolpert's (2018) paper is a very brief research report; he explains that cognate words did not show any facilitation effect because of the presence of false friends in the stimuli list, which led to interference in reading. On the one hand, parallel co-activation of languages was indeed seen during those tasks, via interference from false friends. On the other hand, cognate words seemed not to produce any effect, be it facilitatory or inhibitory. Based on the results from a study by Brenders, van Hell, and Dijkstra (2011, Exp. 3), it would be possible to expect that a stimuli list with both false friends and cognates would lead to inhibition effects for both types of words. However, the participants in Brenders, van Hell, and Dijkstra's study were not minority language speakers and were older than the ones in Woolpert's — respectively, 5th, 6th, 7th, and 9th grades, and 2nd, 3rd, and 4th grades. Thus, the results I report in this study are similar to the ones presented by Woolpert with minority language children.

In the next chapter, I will present Study 2, which had adult speakers of HC living in Brazil as participants.

4 STUDY 2

In this Chapter, I will describe Study 2, which investigated the effect of cross-language activation via cognate words across HC and BP with adult speakers of HC and BP. Study 2 also considered whether reading habits (RH) in BP would influence word recognition and the cognate effect. The impact of RH was included because of studies which suggest that RH may be related with the improvement of language skills in the L1 and in other languages (Camiciottoli, 2001; Artieda, 2007; Sun; Bornstein; Esposito, 2021). Although the evidence from these studies is still weak, this proposed influence is in line with other accounts and the more robust evidence behind them, such as the Lexical Quality Hypothesis (Perfetti, 2007). This chapter is composed of the method used in Study 2, information about participants, instruments and procedures, the results with descriptive and inferential statistical analyses, and the discussion of these results in relation to the hypotheses for the study. Research questions, specific objectives, hypotheses, and rationales are described in Section 4.1.1.

The Haitian Creole speaking population in Brazil arrived in the country migrating from Haiti and the Dominican Republic due to a sociopolitical crisis which is present in Haiti since before the earthquake in 2010. So, in Study 2, I expected to observe how the knowledge of HC as a first language would influence the word recognition skills of adult who had been through the educational process (at least partially) and were making use of BP as a majority language in order to be part of the Brazilian society. However, I never intended to compare performance of participants in Study 1 in relation to Study 2 due to their clear maturational differences and life experiences. In Study 2, there were 16 female participants and 19 male participants (N = 35) whose mean age was 32.14 (age range = 18 to 52, SD = 8.37). They were adult speakers of HC and BP and migrated to Brazil. A more detailed description of participants is offered in Section 4.1.5.

Participants completed four tasks and two questionnaires. These tasks were the following ones:

1. a visual lexical decision (LD) task in BP, with cognate and noncognate words across HC and BP, the same one from Study 1;
2. a visual self-paced sentence comprehension (SC) task in BP, with cognate and noncognate words across HC and BP;
3. an auditory self-paced SC task in BP, with cognate and noncognate words across HC and BP;

4. a reading habits (RH) in BP questionnaire — part of Hübner *et al.* (2019);
5. a HC receptive vocabulary test based on Lima (2007);
6. and a language history and use questionnaire — part of Li *et al.* (2019).

The construction of these tasks is described in detail in Section 4.1.4. The procedures and steps used during data collection sessions are listed in Section 4.1.6. Inferential data analyses examined whether the cognate status of words would have an effect on RTs and accuracy of participants during the written LD task and written and spoken SC tasks and whether RH would have an effect on these variables and/or would influence the cognate effect. A brief summary of the results would be that no significant cognate effect was seen in any task, any task modality or any independent variable. Also, RH in BP had no significant effect on any independent variable nor interacted reliably with cognate status of words. The results and data analyses are reported in detail in Section 4.2.

4.1 METHOD

4.1.1 Study 2: specific research questions, specific objectives, and hypotheses

Study 2 investigated whether adults who are native speakers of HC show effects of cross-linguistic interaction between HC and BP on lexical access during comprehension of spoken and written BP. Study 2 was guided by the following research questions:

1. Do cognate words across HC and BP, in comparison with noncognate words, facilitate visual lexical access in BP?
2. Do cognate words across HC and BP, in comparison with noncognate words, facilitate visual and auditory sentence comprehension in BP?
3. Do reading habits (RH) in BP influence visual lexical access in BP?
4. Do RH in BP influence visual and auditory sentence comprehension in BP?

In order to answer these research questions, Study 2 pursued the following specific objectives:

1. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during a visual lexical decision task in BP;

2. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce higher accuracy rates during a visual lexical decision task in BP;
3. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during visual and auditory sentence comprehension tasks in BP;
4. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during a visual lexical decision task in BP;
5. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce higher accuracy rates during a visual lexical decision task in BP;
6. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during visual and auditory sentence comprehension tasks in BP.

Objectives 1 and 2 are related to research question 1; Objective 3 is related to research question 2; Objectives 4 and 5 are related to research question 3, and Objective 6 is related to research question 4. The Hypotheses for Study 2 are:

1. Cognate words, in comparison with noncognate words, produce shorter RTs during a visual lexical decision task in BP;
2. Cognate words, in comparison with noncognate words, produce higher accuracy rates during a visual lexical decision task in BP;
3. Cognate words, in comparison with noncognate words, produce shorter RTs during both visual and auditory sentence comprehension tasks in BP;
4. More frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during a visual lexical decision task in BP;
5. More frequent RH in BP, in comparison with less frequent RH, produce higher accuracy rates during a visual lexical decision task in BP;
6. More frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during both visual and auditory sentence comprehension tasks in BP.

The rationale for hypotheses 1 and 2 derives from the same evidence that supports the rationale for hypotheses 1 and 2 of Study 1, with the inclusion of studies on sentence comprehension. Jared (2015) explains that the two languages of a bilingual person may interact since they are always activated at some level. This has been observed in word recognition and sentence reading tasks especially with adults. For example, Arêas da Luz Fontes and Schwartz (2015) inserted homonym cognate and noncognate words and non-homonym cognate and non-cognate words in sentences whose contexts were either neutral or biased towards one of the homonyms' meanings. In L2 sentences with biased contexts, co-activation of meanings across languages via cognates seemed to facilitate access to both meanings of the homonym word, especially for participants with high working memory span. In addition, the study by Bultena, Dijkstra, and van Hell (2014) showed that cognate facilitation effects occur in L2 sentence context. However, there are restrictions to its magnitude: they observed smaller effects of cognate words for participants who reported high L2 proficiency. Further, Schwartz and Kroll (2007), Titone *et al*. (2011), Allen, Conklin and Miwa (2020) and Lijewska (2023) observed smaller effects or no effects from cognate words due to sentence constraint. However, van Assche *et al.* (2011) observed similar facilitation effects in low-constraint and in high-constraint sentences. Also, there are studies showing cross-language activation in oral language comprehension with isolated words, word pairs and in sentences (Marian; Spivey, 2003; Thierry; Wu, 2007; van Dijk; Dijkstra; Unsworth, 2022). These results allow us to expect that cognate words would be associated with a facilitation effect which may be weaker in comparison with the ones found in word recognition tasks.

The rationale for hypotheses 3 and 4 is that reading habits influence language development in relation to both L1 and L2. Pratheeba and Krashen (2013) showed that more frequent reading habits are moderately and positively correlated with L2 vocabulary size. Similarly, Santos-Díaz (2017) found that the frequency of reading and the size of L2 active and passive vocabularies were correlated. Rudell and Hu (2010) observed that more exposure to a language leads to higher experience with words, which in turn contributes to higher speed of lexical access. In addition, Artieda (2007) mentioned that reading habits in the L1 can influence the L2 in that L1 reading experience may function as a threshold for developing a L2. She adds that, for advanced learners, reading habits in the L2 may have greater impact than L1 spelling skills. These findings inspire the idea that L2 reading habits of adult speakers of HC will influence the development of their L2 reading skills.

4.1.2 Experimental design

This is a cross-sectional, quasi-experimental study (Christensen; Johnson; Turner, 2015). The independent variable is the cognate status of words, either cognate or noncognate, which means that the variable is categorical and has two levels. The dependent variables are response times (RTs) in milliseconds, a continuous variable in the lexical decision task and in both visual and auditory sentence comprehension tasks, and response accuracy, a binary variable in the lexical decision task. The covariable is RH in BP, a continuous variable measured by the scores in the reading habits test.

4.1.3 Selection of cognate words

The procedures for selecting cognate words across HC and BP for Study 2 were the same ones for Study 1 presented in Section 3.1.3.

4.1.4 Instruments

All tasks in Study 2 can be accessed through the following link: https://antigo-labling.ufsc.br/estudos/pietra/adultos/. The final stimuli list for each task is available on OSF (https://osf.io/c5pzj/files/osfstorage). These tasks are the visual lexical decision task in BP used in Study 1, a visual and an auditory self-paced sentence comprehension task in BP, a RH in BP questionnaire, and a HC vocabulary task. Also 20 questions from the Language History Questionnaire (LHQ 3.0) were selected to be answered by participants. These tasks are described below.

4.1.4.1 The written lexical decision (WLD) task

The written lexical decision task of Study 2 is the same task used in Study 1, presented in Section 3.1.4.1. Participants took an average of 13.31 minutes (*SD* = 6.53) to complete this task.

4.1.4.2 The written self-paced sentence comprehension (WSC) task

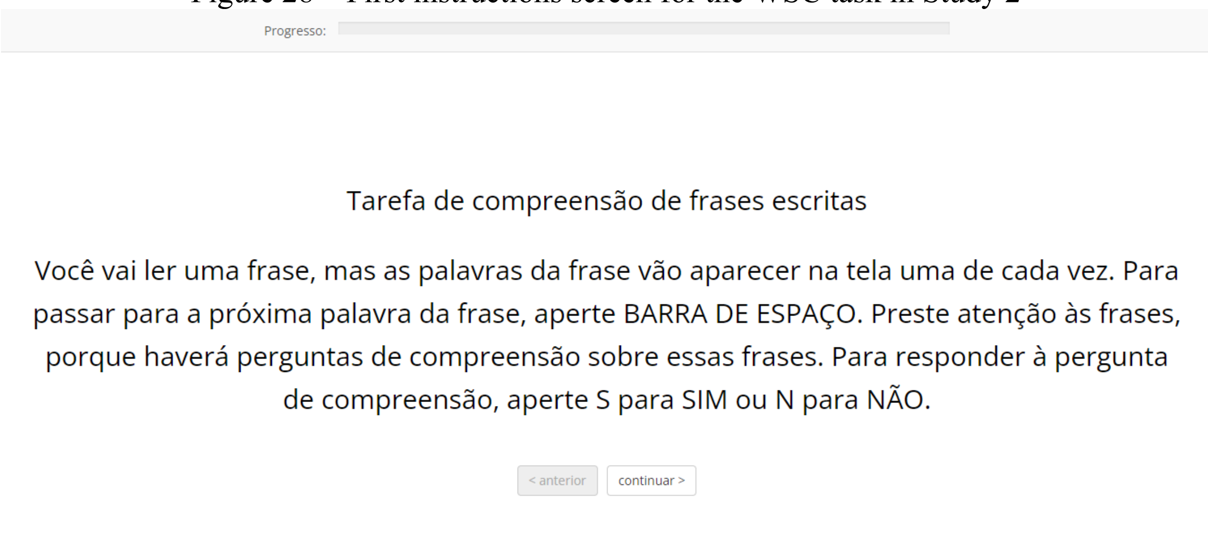
This was a self-paced sentence reading task in BP aimed at testing lexical access in sentence context. The task consisted of 40 sentences in the active voice containing cognate words and 40 sentences in the active voice containing noncognate words as target words. The sentences were created for the purposes of Study 2 and went through two rounds of tests of plausibility in BP. Sentences were evaluated in relation to how natural they sounded in BP.

Feedback was collected about any aspect that made them unnatural (for example, unexpected words for the context of the sentence). Sentences were modified until they were considered plausible. In each round, a total of 8 fellow researchers from LabLing voluntarily filled in an acceptability form. All target words were the fifth word in the sentences. Sentences varied from 6 words to 9 words (*M* = 7.71; *SD* = 0.78). Target words were always the fifth word in the sentences. This way, it would be possible to test the cognate effect on the target words and to verify whether any spillover effect had impacted RTs on the sixth word of the sentences (Duffy; Morris; Rayner, 1988). Target words were selected from the pool of highly frequent ones we created with the help of an informant; this stimuli selection process was described in detail in section 3.3 above. The cognate and noncognate words used in this task were unique and not used in any of the other tasks. It was not possible to control for target word length in a way which would keep it constant. The length of cognate words ranged from 4 to 10 letters (*M* = 6.85; *SD* 1.61), whereas the length of noncognates ranged from 2 to 11 letters (*M* = 6.65; *SD* = 2.11). For each type of word (cognates and noncognates), there were 28 nouns and 12 adjectives. There were no verbs because of the position of target words in the sentence (always the fifth word). Non-target words were selected from the pool of non-cognate words mentioned in section 3.3.

A comprehension question was presented after every 5 sentences to make sure participants were paying attention to the sentences. These questions were *yes* or *no* questions about the information presented in the previous sentence. For example, the sentence *Ele não gosta de tomate misturado na comida.* (He doesn't like to mix tomatoes in his food.) was followed by this comprehension question *Ele gosta de tomate na comida?* (Does he like tomatoes in his food?). Since there were 40 sentences with cognate words and 40 sentences with noncognate words, 80 sentences in total, there were 16 blocks of sentences and 16 comprehension questions. Sentence presentation was divided into blocks of 5, so the last sentence was followed by its respective comprehension question. Because of difficulties in programming full randomization in the task and making sure the comprehension question was presented right after its respective sentence, the order of sentences inside blocks was fixed. However, block presentation was randomized for each participant. It is important to mention a serious error which occurred at this point of task construction and was identified only after data collection: the cognate status was not mixed within sentence blocks as it should have been. This means that all participants were presented blocks of 5 cognate sentences and blocks of 5 noncognate sentences. Thus, a participant might have noticed that there were words similar across languages in those 5 sentences. This flaw would have been enough to question any co-activation effects seen in this task.

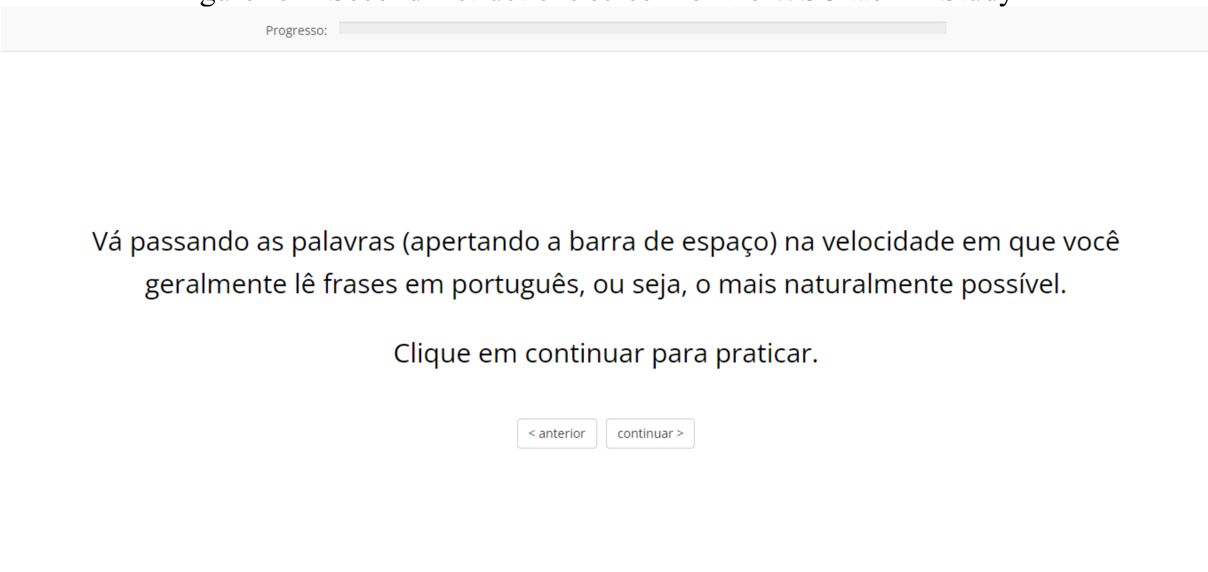
All tasks in Study 2 were completed on an Internet browser on a laptop or desktop computer. For the WC task, participants were instructed to read the sentences at a natural pace, as if they were reading any material in BP, and to pay attention to every sentence because sometimes a comprehension question would appear. Participants were not informed that comprehension questions appeared after every 5 sentences; this instruction was purposefully vague. There was a practice section with two sentences and one question. Participants were encouraged to repeat the practice section if they had doubts or had answered the question wrong. Sentences were presented on screen one word at a time. Participants would press the space bar on the keyboard after reading a word to move to the next one. On comprehension questions, they would press either S for yes and N for no. Words appeared in the center of the screen, in black Arial Sans Serif font, size 24 point. Figures 28 to 37 illustrate the instructions screens and how a sentence trial was displayed.

Figure 28 – First instructions screen for the WSC task in Study 2



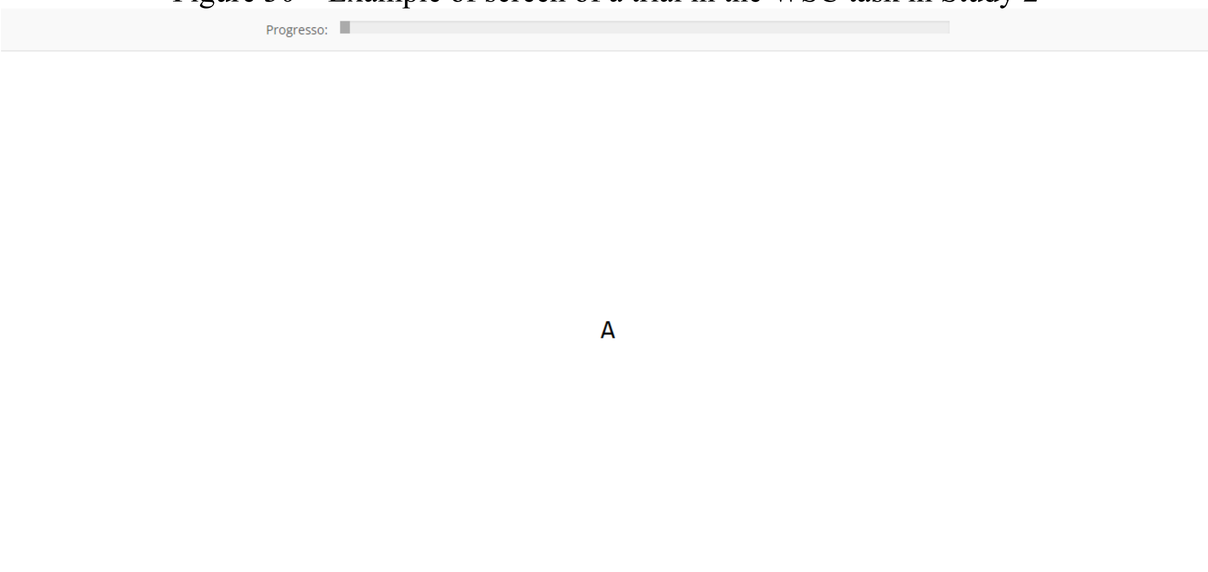
Source: The Author (2024)

Figure 29 – Second instructions screen for the WSC task in Study 2



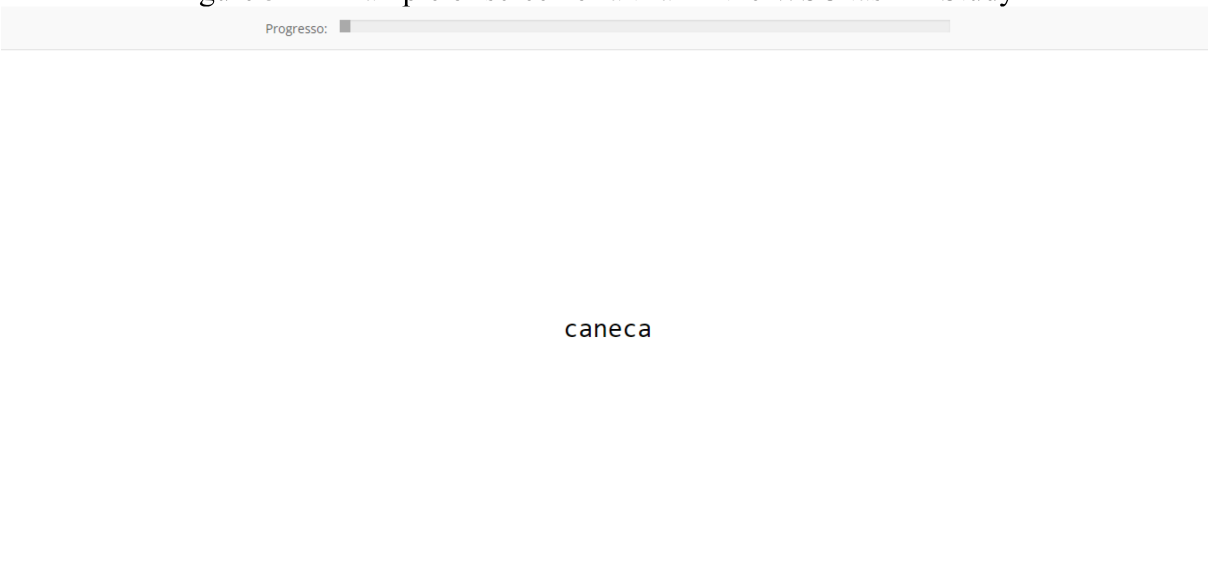
Source: The Author (2024)

Figure 30 – Example of screen of a trial in the WSC task in Study 2



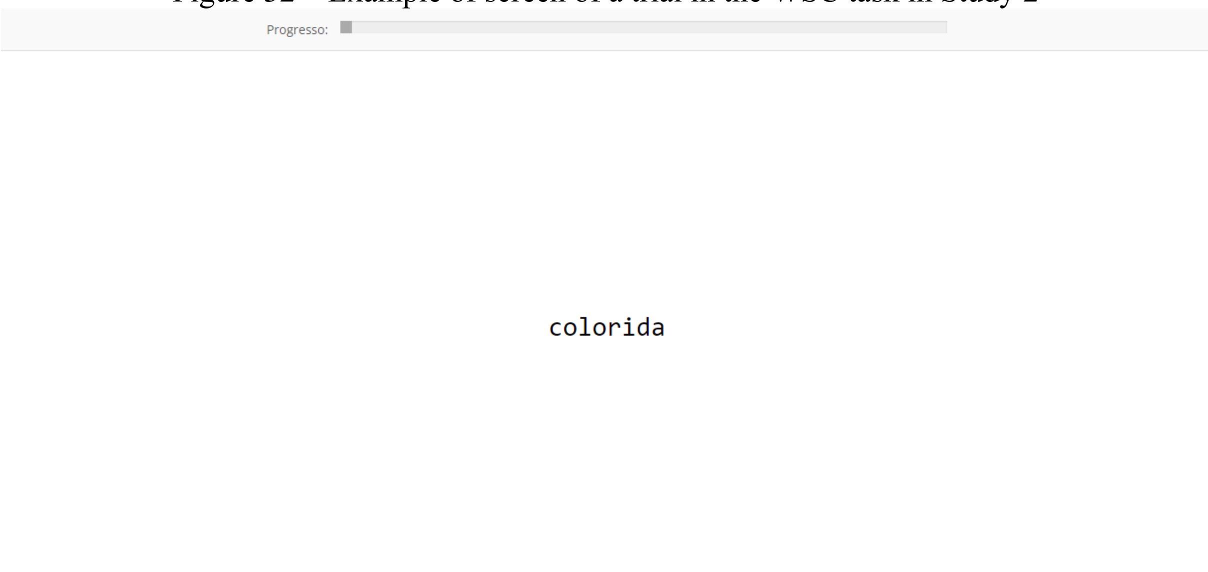
Source: The Author (2024)

Figure 31 – Example of screen of a trial in the WSC task in Study 2



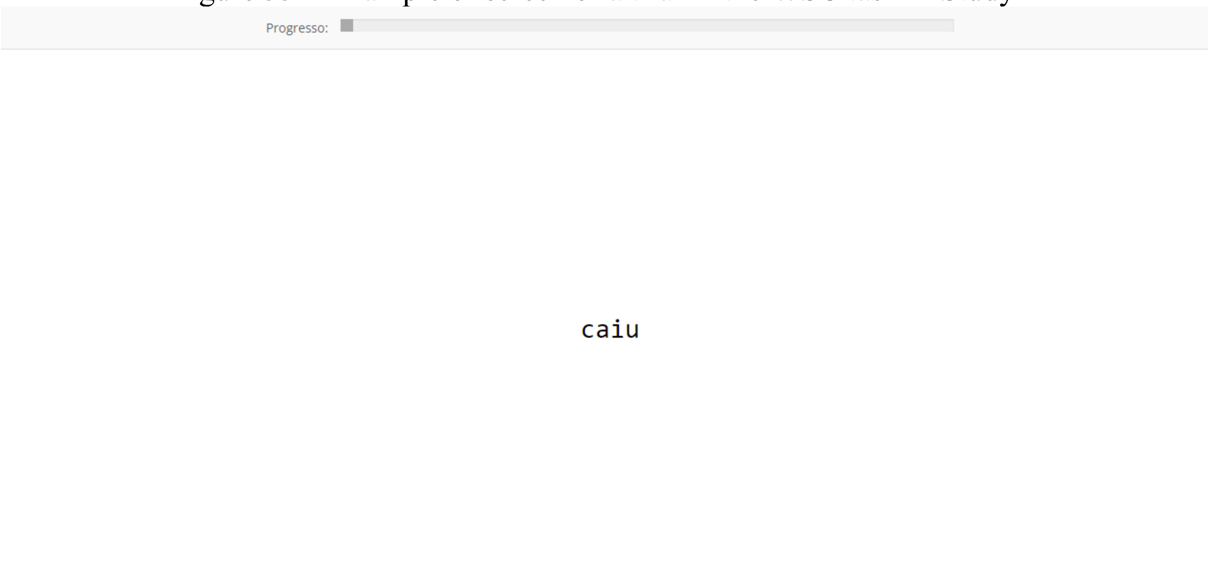
Source: The Author (2024)

Figure 32 – Example of screen of a trial in the WSC task in Study 2



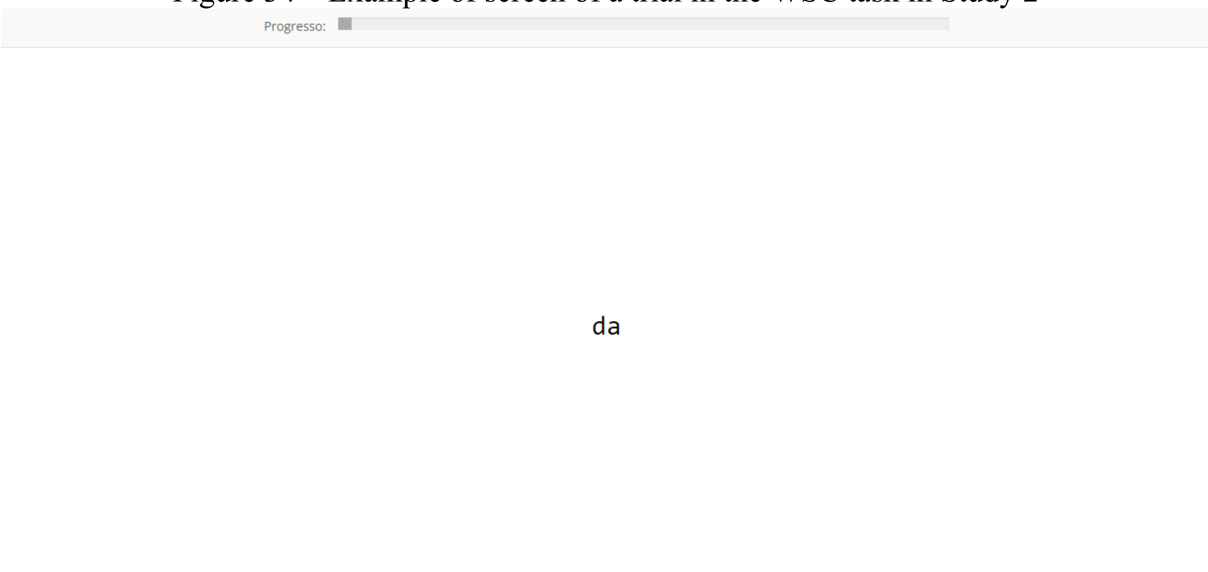
Source: The Author (2024)

Figure 33 – Example of screen of a trial in the WSC task in Study 2



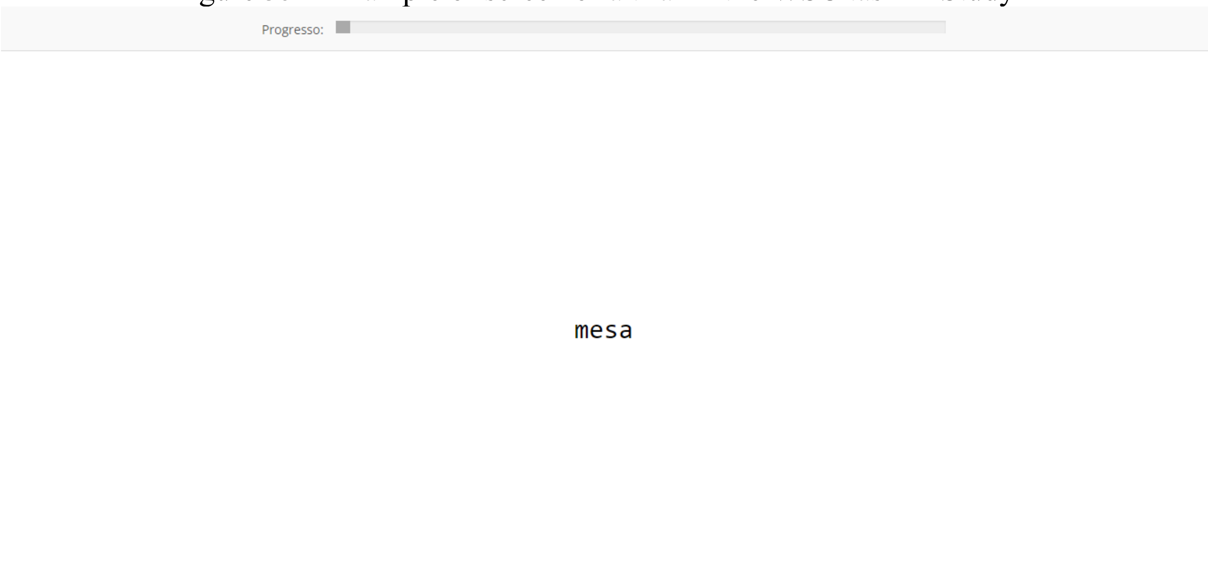
Source: The Author (2024)

Figure 34 – Example of screen of a trial in the WSC task in Study 2



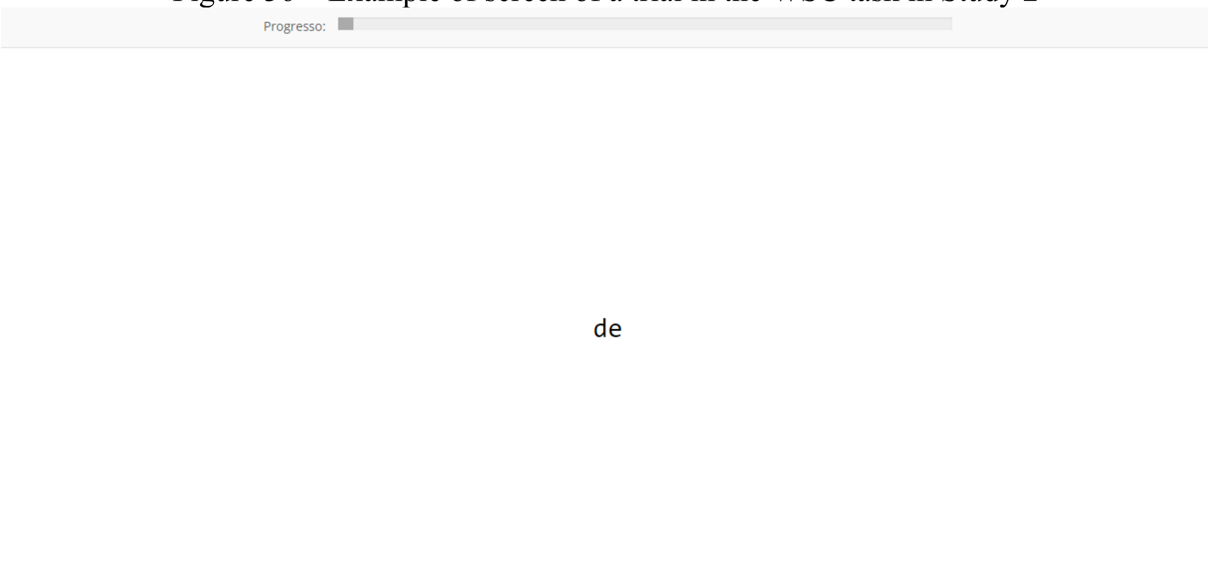
Source: The Author (2024)

Figure 35 – Example of screen of a trial in the WSC task in Study 2



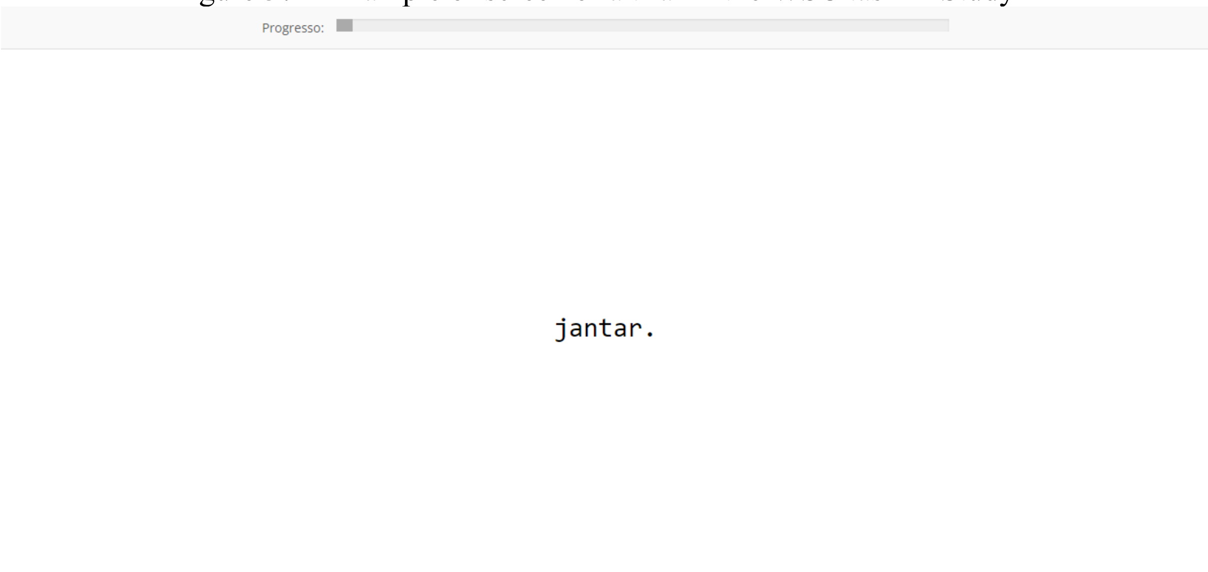
Source: The Author (2024)

Figure 36 – Example of screen of a trial in the WSC task in Study 2



Source: The Author (2024)

Figure 37 – Example of screen of a trial in the WSC task in Study 2



Source: The Author (2024)

Reaction times were registered from stimuli presentation onset and for each word in the sentence as a measure of processing time. There was no time limit for responses. Accuracy was only registered for comprehension questions. Participants took an average of 25.86 minutes (*SD* = 13.87) to complete this task. The final stimuli list used in this task can be examined directly in OSF (https://osf.io/592hm).

4.1.4.3 The spoken self-paced sentence comprehension (SSC) task

This was a self-paced sentence listening task in BP. Task structure characteristics and programming were the same as in the written task described in section 3.4.2.2. The major difference between them was that all stimuli were presented auditorily in this task. Word selection and sentence creation followed the same criteria as in the WSC task. In addition, all experimental words were recorded by two professional voice actors, a man and a woman, contacted via the *Vinte Pila* platform. This means that every sentence word in the task had a respective audio file. In order to counterbalance voice gender when presenting the stimuli, the task was programmed to randomly select one of the two voices when presenting a sentence. In other words, some participants may have heard a sentence spoken by a male voice, and other participants may have heard it spoken by a female voice.

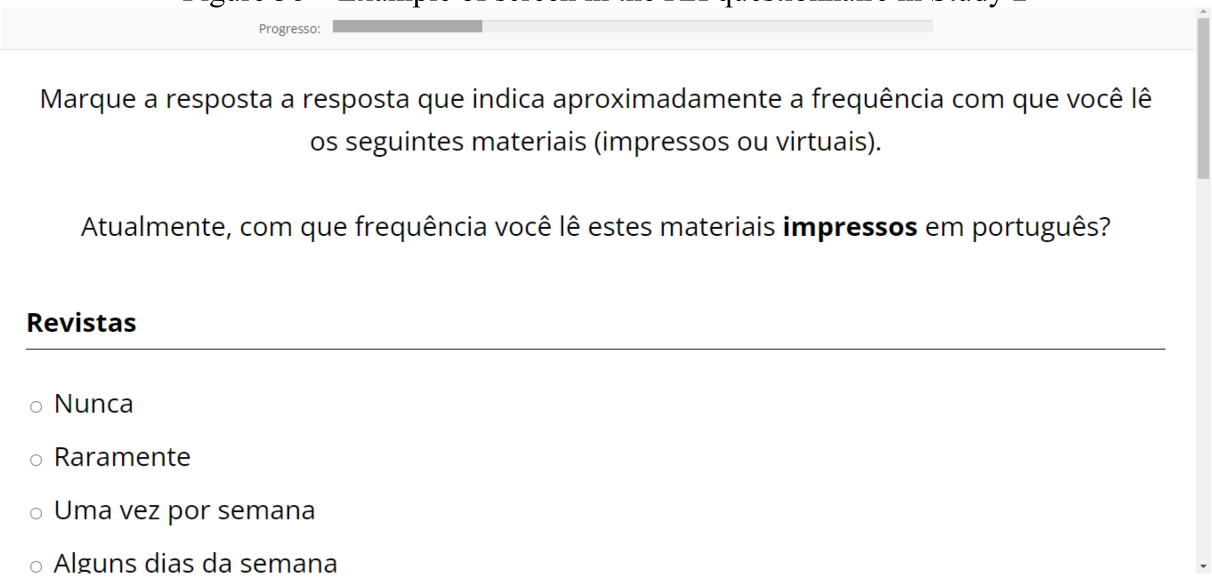
There were 40 sentences with cognate words as targets and 40 sentences with noncognates, 80 sentences in total. A comprehension question was presented after every 5 sentences to make sure participants were paying attention to the sentences. These questions were yes or no questions about the information presented in the previous sentence and written on the screen, not spoken. In total, there were 16 blocks of sentences and 16 comprehension questions. Sentence presentation was divided into blocks of 5, so the last sentence was followed by its respective comprehension question. Just like in the WSC task, the order of sentences inside blocks was fixed — and cognate status was not mixed within blocks. Block presentation was randomized for each participant. Sentences varied from 6 words to 9 words (M = 7.69, SD = 0.67). The length of cognate words ranged from 4 to 11 letters (M = 6.8, SD = 1.89), whereas the length of noncognates ranged from 4 to 10 (M = 7, SD = 1.72). For cognate words, there were 28 nouns and 12 adjectives; for noncognates, there were 29 nouns and 11 adjectives. There were no verbs because of the position of target words in the sentence (always the fifth word). The cognate and noncognate words used in this task were also unique and not used in any of the other tasks.

Instructions were written on the screen in the beginning of the task. Participants were instructed to listen to the sentences at a natural pace, as if they were listening to a conversation in BP, and to pay attention to each sentence because sometimes a comprehension question would appear. There was a practice section with two sentences and one question. Participants were encouraged to repeat the practice section if they had doubts or had answered the question wrong. Sentences were presented one word at a time. Participants would press the space bar on the keyboard after listening to a word to move to the next one. Comprehension questions were written on the screen — so participants would be able to reread the question — and were displayed until an answer was given, and participants would press either S for yes and N for no. Participants took an average of 23.17 minutes (SD = 4.42) to complete this task. The final stimuli list and audios used in this task can be examined directly in OSF (respectively, https://osf.io/7bdup, https://osf.io/4hpaf, and https://osf.io/u2mf7).

4.1.4.4 The Reading Habits (RH) questionnaire

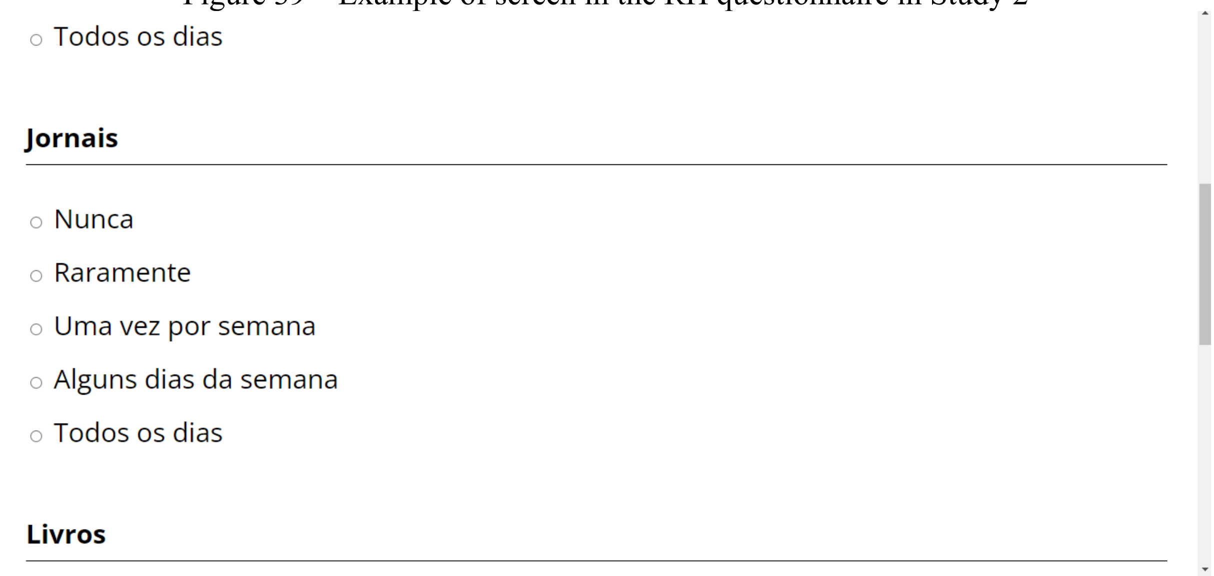
For the purposes of Study 2, this questionnaire was adapted from the reading and writing habits form (Hübner *et al.*, 2019) which is part of the Aging and Language Evaluation Battery (Zimmermann; Deleaere; Fonseca, 2019). Thus, it is part of a validated neuropsychological battery. This questionnaire was selected as a subjective measure of the frequency in which participants read materials in BP. There were 7 very brief questions, and answers were given on a 5-point scale in which zero equaled “never” and 4 equaled “every day”. Questions inquired whether participants read magazines, newspapers, and books, either in paper or virtually, in present days and in the past, and whether they used social media in BP. I selected questions regarding RH in the present, not in the past. Participants took an average of 5.48 minutes (*SD* = 3.23) to complete this questionnaire. The final set of questions used in this questionnaire is in Appendix G. Figures 38 to 41 illustrate how the questionnaire was implemented for data collection.

Figure 38 – Example of screen in the RH questionnaire in Study 2



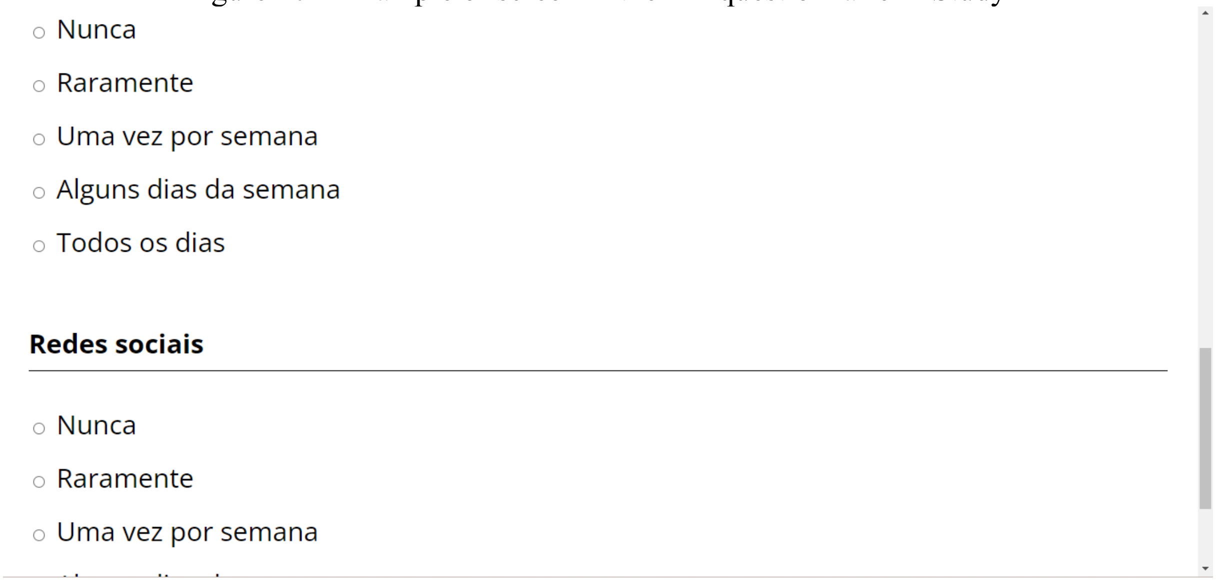
Source: The Author (2024)

Figure 39 – Example of screen in the RH questionnaire in Study 2



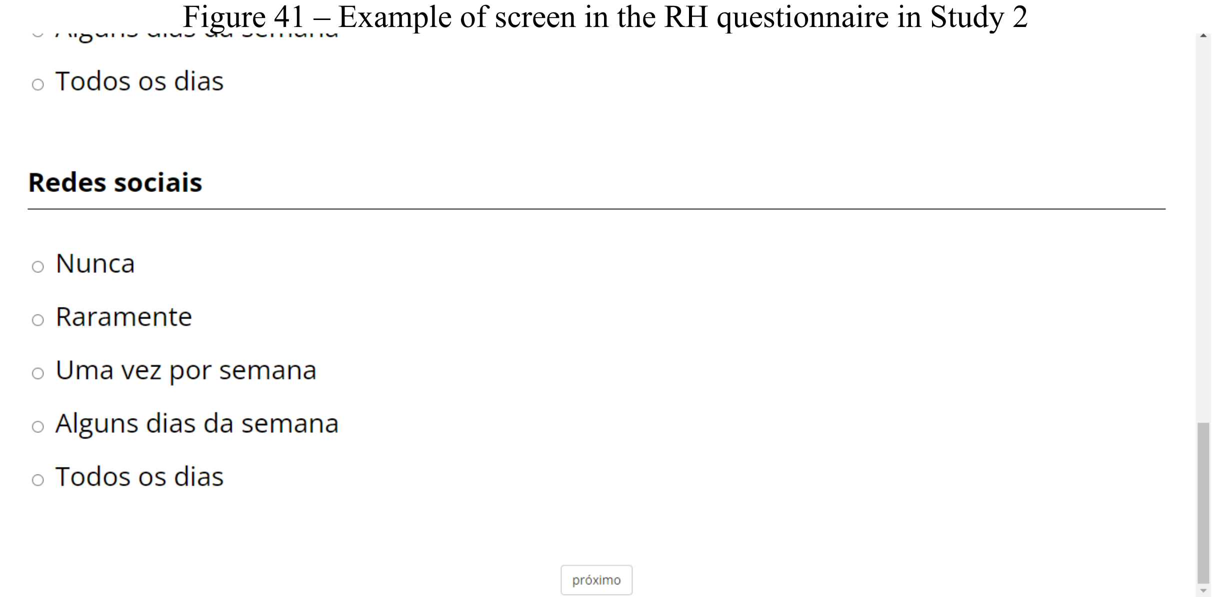
Source: The Author (2024)

Figure 40 – Example of screen in the RH questionnaire in Study 2



Source: The Author (2024)

Figure 41 – Example of screen in the RH questionnaire in Study 2



Source: The Author (2024)

4.1.4.5 The HC receptive vocabulary test

The HC receptive vocabulary test of Study 2 is the same test used in Study 1, presented in Section 3.1.4.5. Participants took an average of 7.56 minutes (*SD* = 2.38) to complete this test.

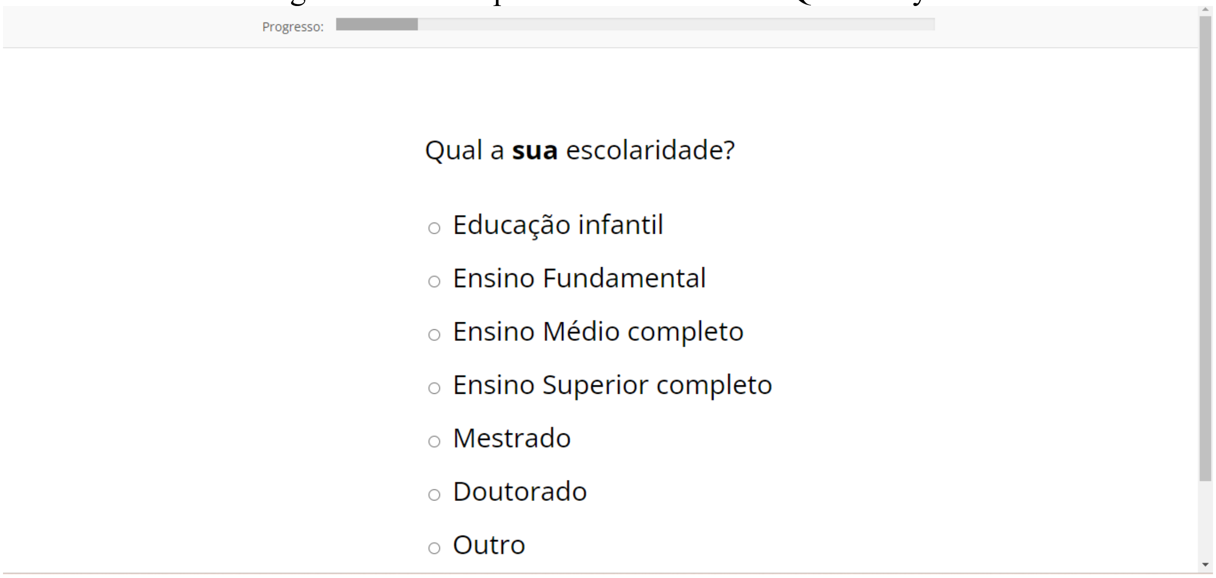
4.1.4.6 The Language History Questionnaire

The LHQ of Study 2 is the same set of questions from LHQ 3.0 used in Study 1, presented in Section 3.1.4.6 (Appendix A). Participants took an average of 45.59 minutes (*SD* = 19.13) to complete this questionnaire. Figures 42 to 46 illustrate how the questionnaire was implemented for data collection.

Figure 42 – Example of screen of the LHQ in Study 2

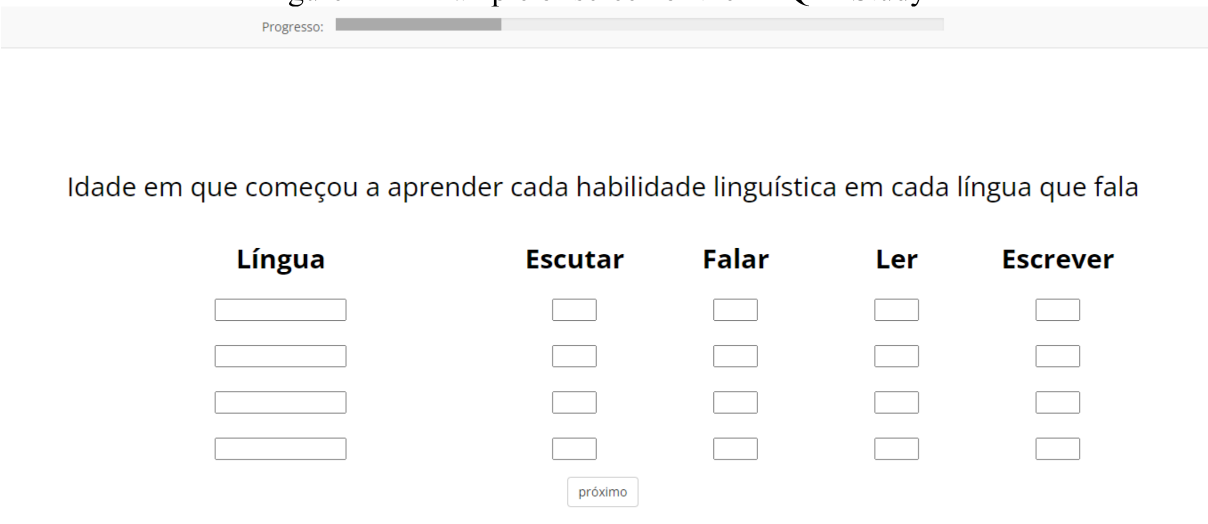
Source: The Author (2024)

Figure 43 – Example of screen of the LHQ in Study 2



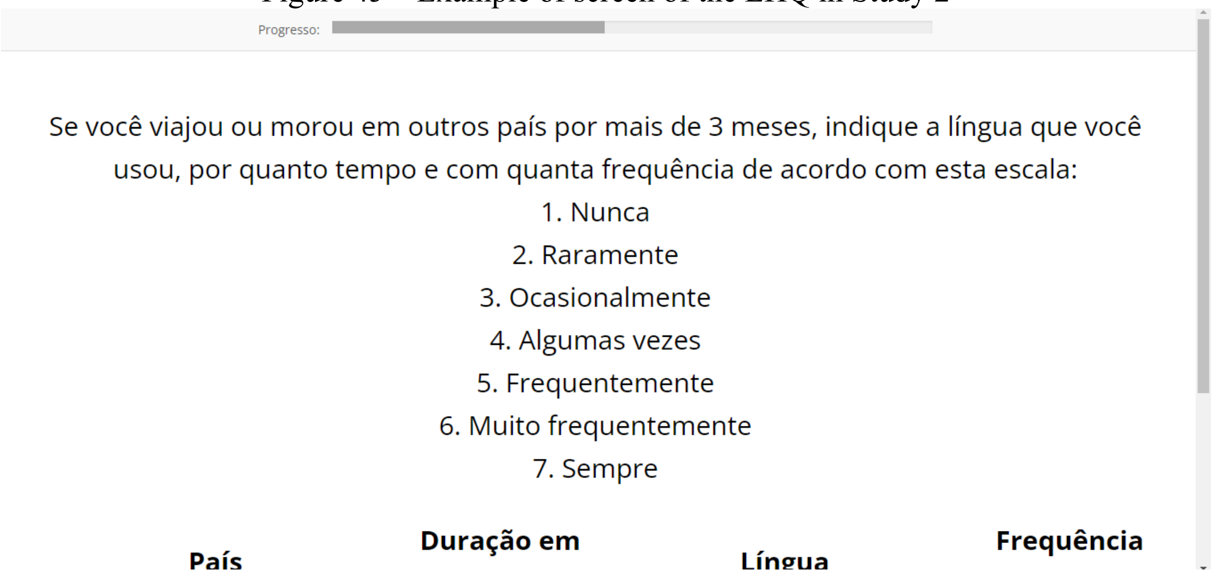
Source: The Author (2024)

Figure 44 – Example of screen of the LHQ in Study 2



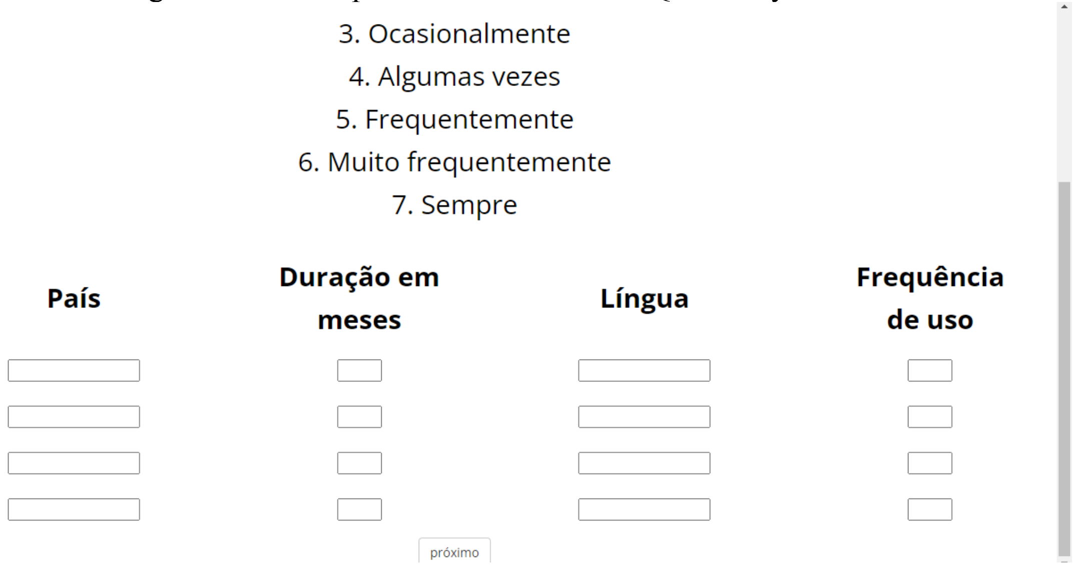
Source: The Author (2024)

Figure 45 – Example of screen of the LHQ in Study 2



Source: The Author (2024)

Figure 46 – Example of screen of the LHQ in Study 2



Source: The Author (2024)

4.1.5 Participants

Participants for Study 2 were HC-speaking adults. Since tasks were programmed in JavaScript and presented on a website in case data collection had to be held remotely, participants could be invited online. An invitation poster was designed in order to invite participants (Appendix H). It explicitly invited people who were speakers of HC and offered two ways (e-mail and phone number) to get in contact with the researcher for anyone interested in participating.

More than 20 adult HC-speakers were invited through mutual acquaintances and at the university; thirty-eight HC-speaker adults were invited at the church. However, only 44 adults participated. Then, the final sample in Study 2 consisted of 35 HC-speaker adults (19 male, 16 female, age range = 18 to 52, $M = 32.14$, $SD = 8.37$). Data from 7 participants was not added because they were too distracted during data collection; data from 2 participants was lost; data from one participant was not added because the entire set of tasks was not completed.

In order to describe the population sample who participated in Study 2, I will refer to answers given to the 20 questions from the Language History Questionnaire (LHQ) presented in Appendix A. Descriptive statistics about participants' language history and use are presented next. These statistics were only calculated for questionnaire items which had answers for at least half of the participants.

Languages spoken, other than HC and BP, were French, Spanish and English. Other than HC, 8 participants reported also speaking French, 8 French and Spanish, 7 French and

English, and 1 French, Spanish and English. Mean age of starting to be exposed to HC was 1.53 (*SD* = 0.86, range = 1 to 4) and to BP was 27.05 (*SD* = 7.77, range = 10 to 47). Considering that some participants did not provide answers to some questions, 19 participants reported using HC more with their family, 6 reported using BP more with their family, 1 reported using both languages equally; 10 reported using HC more with friends, 10 reported using BP more with friends, 5 reported using both languages equally. I insisted with the 35 participants to gather information about the year they arrived in Brazil. The median year of arrival was 2017 and the mean was 2017,029. Table 35 presents the mean age (SDs in parenthesis) when participants started using each language in each environment.

Table 35 – Mean age when participants started using each language in each environment in Study 2

(N = 25)	Home	Friends	School	Work	Computer	Online games
Haitian Creole	1.87 (1.12) Range: 1-5	4,00 (2.22) Range: 1-12	4.68 (2.46) Range: 2-11	17.24 (5.40) Range: 10-25	16,00 (6.32) Range: 10-24	9.88 (8.49) Range: 9-22
Brazilian Portuguese	26.70 (9.39) Range: 5-47	26.40 (9.35) Range: 5-47	27.17 (9.88) Range: 5-47	28.67 (8.70) Range: 5-47	25.60 (10.78) Range: 5-51	26.06 (11.75) Range: 5-51

Source: The Author (2024)

Table 36 presents means and SDs of self-reported proficiency levels for language skills in HC and BP. The scale had 7 points, where was the highest proficiency possible.

Table 36 – Mean (SD) self-reported proficiency levels for language skills in HC and BP in Study 2

(N = 26)	Listening	Speaking	Reading	Writing	All skills
Haitian Creole	6.46 (1.03) Range: 3-7	6.52 (0.82) Range: 4-7	6.46 (0.86) Range: 4-7	6.50 (0.86) Range: 4-7	6.49 (0.09) Range: 6.46-6.52
Brazilian Portuguese	5.35 (1.02) Range: 4-7	4.96 (1.56) Range: 2-7	5.46 (1.24) Range: 3-7	4.96 (1.37) Range: 2-7	5.18 (0.23) Range: 4.96-5.46

Source: The Author (2024)

Table 37 presents means (SDs in parenthesis) for hours per day using a specific language for each activity.

Table 37 – Mean hours per day using a specific language for each activity in Study 2

(N = 26)	Watching TV	Reading for fun	Reading for work	Writing	Using social media
Haitian Creole	3.56 (3.77) Range: 1-16	4.89 (5.05) Range: 0.5-16	3.93 (5.01) Range: 1-16	2.60 (2.56) Range: 0.5-7	4.33 (4.53) Range: 1-18
Brazilian Portuguese	4.33 (4.50) Range: 0.5-18	4.67 (5.09) Range: 1-18	6.29 (4.68) Range: 0.5-18	5.17 (5.14) Range: 1-18	4.43 (4.57) Range: 1-18

Source: The Author (2024)

Participants' and their parents' educational level are reported in Table 38. Some data is missing because some participants did not know their parents' educational level or never sent their answers back.

Table 38 – Number of participants who reported parents' and their own educational level in Study 2

	Preschool	Elementary school	High school	Undergraduate course	Total of answers
Participant	0	6	16	6	28
Mother	5	10	7	2	24
Father	2	12	5	4	23

Source: The Author (2024)

Study 2 was approved by the Ethics Committee at UFSC under CAAE 51539821.3.0000.0121. To take part in Study 2, participants had to read the Consent Form (Termo de Consentimento Livre e Esclarecido, TCLE) while the researcher read it out loud. Any questions were answered, and data collection only started if participants had agreed to participate. The TCLE had versions in BP and in HC (Appendices I and J).

4.1.6 Procedures

Data collection sessions occurred in diverse settings according to space availability and participant convenience: library, university computer room, a quiet café near their workplace, study rooms at UFSC, their home, remotely. The reason for this low control of data collection environment is the fact that participants' work schedules were their priority. Thus, I had to provide them with more convenient places for participating in Study 2. From the 35 participants whose data was included in analyses, 19 participated in a computer room at UFSC; 4 participated remotely; 4 participated at a library; 3 participated at study rooms at UFSC; 3 participated at quiet cafés near their workplaces; 1 participated at home because there was no laptop or computer available and there was no possibility of commuting to the university.

The 19 participants who took part in Study 2 at a computer room at UFSC were invited at a church in Palhoça. I describe this invitation process in detail in Chapter 6. This group of 19 participants only had Saturdays available for taking part in the present study and had no laptop or computers available for remote participation. Thus, I scheduled data collection sessions at a computer room at UFSC on two Saturdays in September 2022 and I booked a van to drive participants from their church to UFSC and back to their church on both days. The funding for driving these participants to UFSC was taken from the taxa de bancada of the CNPq PQ grant of my supervisor (process number 311632/2019-0). The first data collection session at the computer room lasted 3h30 because there were delays in transportation; the second one lasted 2h30. Due to the duration of the participation sessions, participants were offered cookies, juice and water on both Saturdays. My supervisor and I paid for these snacks, respectively, with funding from the taxa de bancada mentioned above and with personal money.

Next, I will describe the data collection steps used in every participation session, independently of where it was being conducted. Participants were welcomed in the room or in the online video call room, sat in front of the laptop or desktop computer and were presented with the website where the experimental tasks were hosted. It is necessary to mention that another limitation of the present study is that no question regarding machine specifications was added to the experimental task during programming. Thus, I do not have information about the machines used by participants who completed tasks entirely remotely. Tasks were completed either on laptops or desktop computers depending on the machines available. I always took my personal laptop in data collection sessions (DELL Inspiron 3583 15-inch screen, Intel Core i5-8265U, 1.60GHz/1.80 GHz, Windows 11, refresh rate 60Hz). My colleague who usually voluntarily assisted during data collection also took her personal laptop (MacBook Air 13-inch screen 2017, Intel Core i5 1.8 GHz Dual-Core, Intel HD Graphics 6000 1536 MB, refresh rate 60Hz). Desktop computers available in the computer room at CCE UFSC were similar to the following model: Intel Core i3-9100 3.60GHz, Windows 11, DELL 15-inch screen monitor, refresh rate 60Hz. No Bluetooth keyboard was used; only built-in and chord keyboard were used. Participants were motivated to use headphones or earphones during data collection. Some participants preferred to use their own, but most of them used the following models: Philips Bluetooth headphones TAUH202WT/00, 20-20,000 Hz, 102 dB and Inova FON-10066 earphones. It was not possible to control for participants distance from the screen. Tasks were programmed to be accessible via web browser. In most cases, the web browser used was Google Chrome. However, in the few cases when Chrome was not working, Mozilla Firefox was used.

Data collection instruction for the experimenter are described in the data collection notes file (Appendix L). First, participants read the consent form (Appendices I and J) while the researcher read it out loud and asked questions if needed — in person or via WhatsApp. Participation only began after they had agreed to taking part in the Study 2. Participants were reminded that they were free to give up participation at any time with no negative consequences for them. The order in which participants would complete the tasks was prepared in advance and counterbalanced as shown below.

1. Written lexical decision
2. Spoken self-paced sentence comprehension task
3. Written self-paced sentence comprehension task
4. Reading habits questionnaire
5. HC vocabulary test
6. LHQ

Or

1. Written self-paced sentence comprehension task
2. Spoken self-paced sentence comprehension task
3. Written lexical decision
4. HC vocabulary test
5. Reading habits questionnaire
6. LHQ

The priority was presenting either the sentence comprehension tasks or the lexical decision first to make sure participants would not be tired from the tasks. The spoken sentence task was always between two written tasks. Then, participants would complete either the vocabulary or the reading habits test. Finally, the LHQ was presented so that participants could talk to the researcher if they had any problem understanding a question.

Participants were given breaks whenever they needed to use the toilet or to drink water. In those cases, we asked them to finish the task first. All participants wore earphones during the auditory tasks.

4.1.7 Data analysis plan

Data was analyzed using the packages lme4 (Bates *et al.*, 2015) and lmerTest (Kuznetsova; Brockhoff; Christensen, 2017) for mixed-effects linear models in RStudio (Posit Team, 2023). Items and participants were added as random effects. Both random intercepts and random slopes were added at first. If models did not converge, only random intercepts were kept. Six models were run for the planned analyses so that analyses for visual sentence comprehension and spoken sentence comprehension tasks were kept separate.

All planned models had cognate status and reading habits in BP added as fixed effects. Models 1 and 2 analyzed data from the LD task. Model 1 had RTs as the dependent variable for the LD task, and model 2 had accuracy rates. Models 3 and 4 analyzed data from the WSC task. Model 3 had RTs as the dependent variable for the 5^{th} word in the sentences, which was the target word, always either cognate or noncognate. Model 4 had RTs as the dependent variable for the 6^{th} word in the sentences to examine whether the fact that the previous word was either cognate or noncognate would had any spillover effect on the RTs for the next word in the sentence. Models 5 and 6 analyzed data from the SSC task. Model 5 had RTs as the dependent variable for the 5^{th} word in the sentences while model 6 had RTs as the dependent variable for the 6^{th} word in the sentences.

Significant interactions were analyzed by running two more mixed-effect linear models, and p-values were corrected by using the function p.adjust() from the stats package (R Core Team, 2023). A pilot analysis was run on the data from 4 participants.

The results section will present first the planned analyses followed by the exploratory analyses.

4.2 RESULTS

All steps taken during data compilation, cleaning and trimming and during data analysis can be examined in the Rmarkdown file available at the project page for this study on OSF (https://osf.io/73tru). First, I will present descriptive statistics and explain the data trimming process. Then, I will present the inferential analyses and the goal for each test. Finally, I will present exploratory analyses which were considered relevant during planned analyses.

4.2.1 Descriptive statistics analyses

Before data trimming, descriptive statistics were calculated. Table 39 shows mean accuracy scores, standard deviations, ranges, and total scores for each task.

Table 39 – Mean accuracy scores, standard deviations, ranges, and total scores in Study 2

Task	Mean accuracy scores	SDs	Range	Total
Written Lexical Decision task	163.3	31.1	120—224	240
Spoken Sentence Comprehension task questions	-	-	-	16
Written Sentence Comprehension task questions	10.9	3.2	5—16	16
HC Vocabulary test	117.9	90.7	76—130	130

Source: The Author (2024)

Table 39 does not present mean accuracy scores for the SSC task because responses to the comprehension questions failed to be recorded for this task. Table 39 indicates that there were participants who presented error rates higher than 30% in the LD task. A score of 168 meant an accuracy rate of 70%. This information was used to select data to include in the inferential analysis. Moreover, scores on the vocabulary test were used to filter data, which will be described below. Table 40 presents mean response times, standard deviations, and ranges for the LD and the SC tasks independently of word position.

Table 40 – Mean RTs, SDs, and ranges for LD and SC tasks in Study 2

Task	Mean RTs	SDs	Range
Written Lexical Decision task	7530.594	459251.600	0—42087083
Spoken Sentence Comprehension task	628.448	14613.160	-1522—1827355
Written Sentence Comprehension task	5699.174	566842.900	1—84233665

Source: The Author (2024)

Table 40 demonstrates that RTs varied vastly which may be observed on SDs and range. Considering these characteristics of the data, some trimming was needed. At this point, there were 8,400 observations in the LD task file, 21,525 in the SSC task file, and 22,085 in the WSC task file.

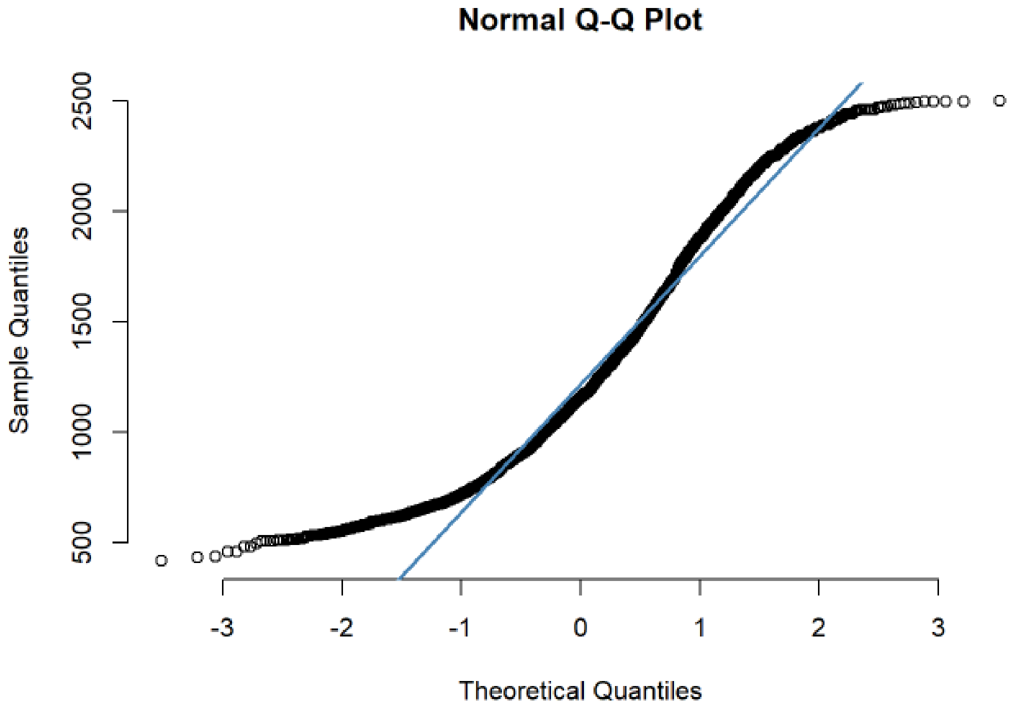
Steps for trimming data were the following. First, I used a threshold of 30% of error rates in the LD task for removing participant data from further analyses (Arêas da Luz Fontes; Schwartz, 2015). This means that only data from participants who had accuracy scores equal or higher to 70% in the LD task were included in inferential analyses for the LD task. At this point, there were 3,120 observations in the LD task file. In addition, I used a threshold of 50% of error rates in the comprehension questions in the WSC task for removing participant data from further analyses. This means that only data from participants who had accuracy scores equal or higher to 50% in the comprehension questions of the WSC task were included in inferential analyses for the SC tasks; this lenient threshold was used because accuracy rates in the comprehension questions for the SC tasks may have been affected by participants' word knowledge. During data collection, some participants mentioned that they did not know the meaning of some words in the comprehension questions. Since the goal of these questions was to make sure participants were paying attention to the sentences being presented, I did not compile data from participants who demonstrated to not be attentive during data collection, that is, participants who were engaging in conversation or were repeatedly looking away from the screen during the task. At this point, there were 19,680 observations in the SSC task file and 20,192 in the WSC task file.

Then, I calculated mean RTs for each participant and removed the ones which were beyond -2.5 SDs or + 2.5SDs from participant mean (Beatty-Martínez *et al.*, 2020). The number of observations was 3120 in the LD task file. There was no change in the number of observations for the SSC and the WSC task file. In addition, I removed RTs shorter than 300 ms and longer than 2,500 ms (Beatty-Martínez *et al.*, 2020). Now there were 2,272 observations in the LD task file, 8,667 in the SSC task file and 15,490 in the WSC task file.

Finally, I removed data from participants who reported knowing less than 50% of words in the HC vocabulary test. At this point, there were 2,272 observations in the LD task file, 8,667 in the SSC task file, and 15,490 in the WSC task file. This indicates that 72.95% of the data in the LD task file, 59.73% in the SSC task file and 29.86% in the WSC task file were removed from further analyses.

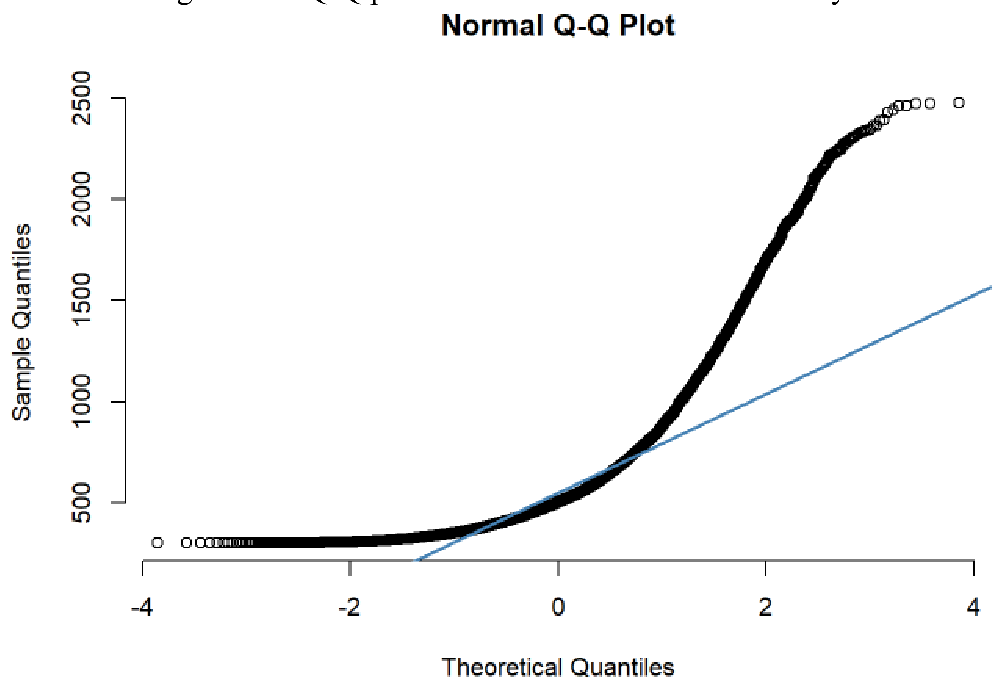
After trimming, the distribution of RTs was examined for normality using the Anderson-Darling normality test (package nortest by Gross and Ligges (2015)) because of the number of observations in the WSC task data file (> 5,000). Response times in all tasks were not normally distributed (LD task: $A = 40.114$, $p < 0.001$; SSC task: $A = 561.4$, $p < 0.001$; WSC task: $A = 636.11$, $p < 0.001$). This was also demonstrated with Q-Q plots, which may be examined in Figures 47 to 49, and with histogram plots, which may be observed in Figures 50 to 52.

Figure 47 – Q-Q plot for RTs in the LD task in Study 2



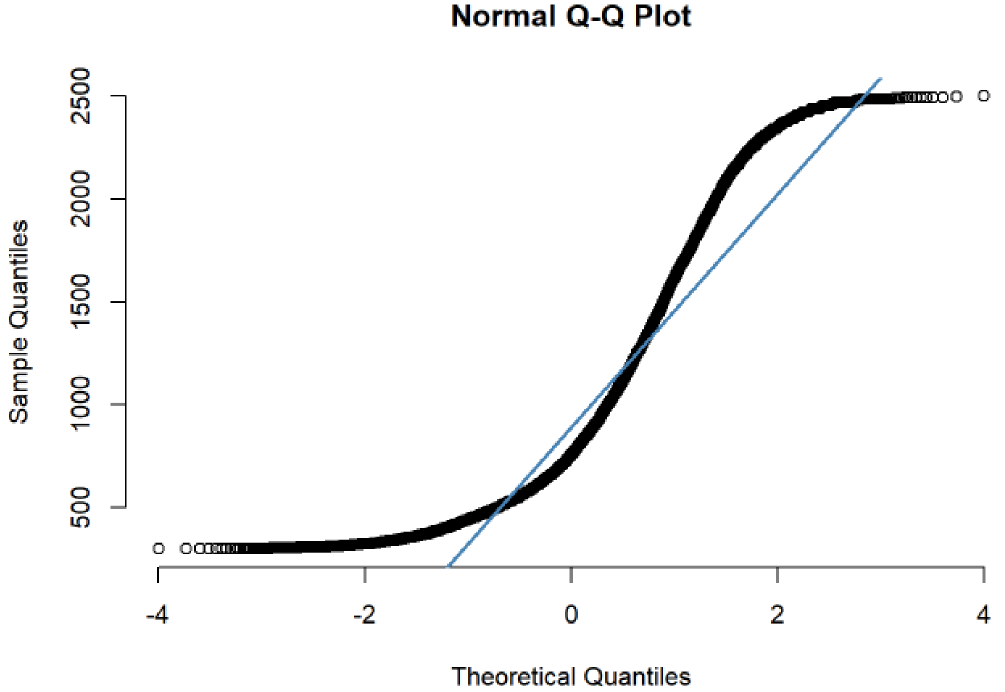
Source: The Author (2024)

Figure 48 – Q-Q plot for RTs in the SSC task in Study 2



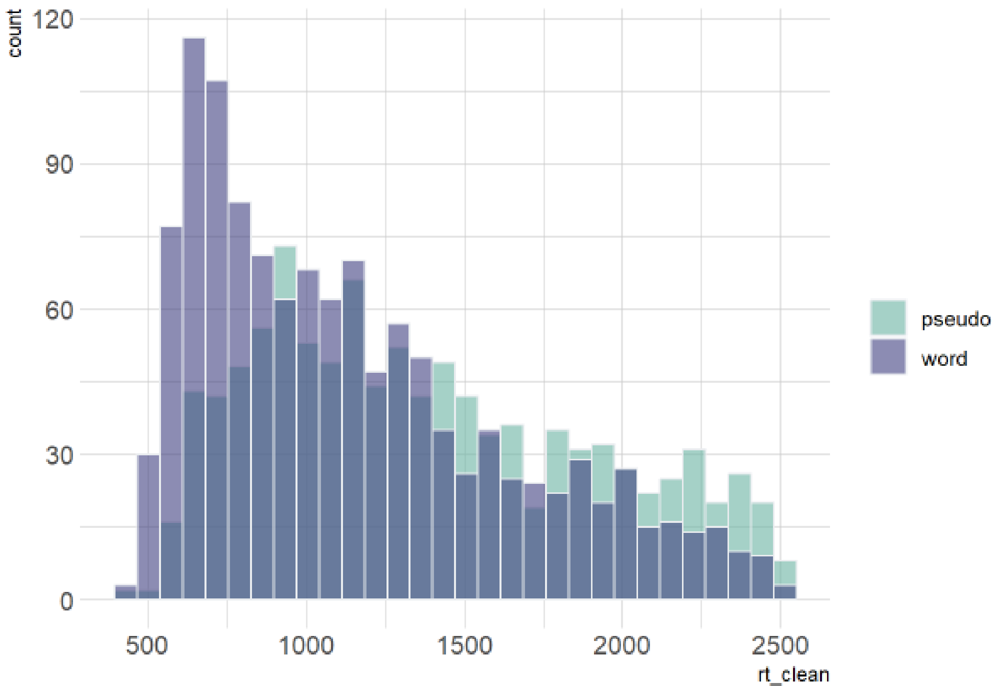
Source: The Author (2024)

Figure 49 – Q-Q plot for RTs in the WSC task in Study 2



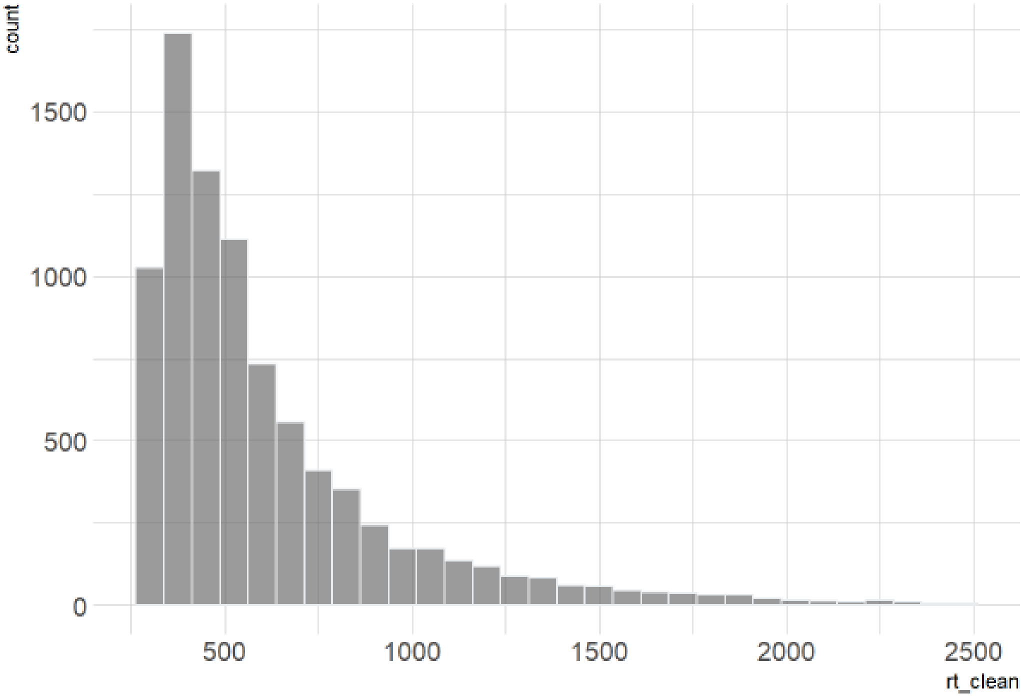
Source: The Author (2024)

Figure 50 – Histogram plot for RTs in the LD task in Study 2



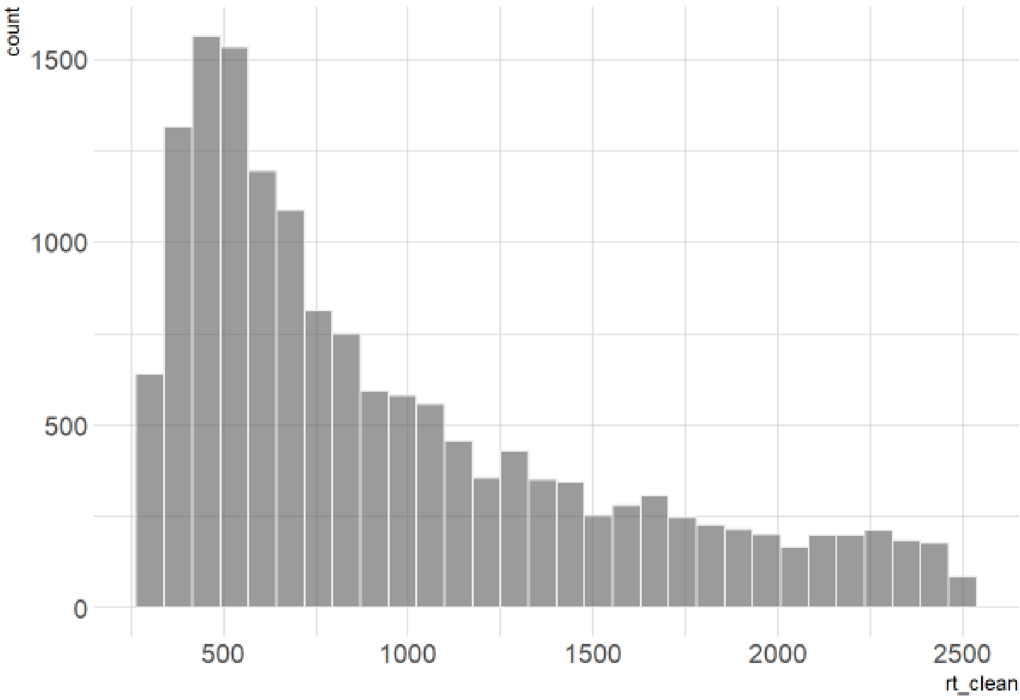
Source: The Author (2024)

Figure 51 – Histogram plot for RTs in the SSC task in Study 2



Source: The Author (2024)

Figure 52 – Histogram plot for RTs in the WSC task in Study 2



Source: The Author (2024)

4.2.2 Inferential statistics analyses

The goal of the lexical decision task is to decide whether a string of letters (or a sequence of phonemes) is a real word or not. One of the effects seen in this task is that real words are responded to faster than pseudowords. Since the distribution of RTs was not normal, I used a Wilcoxon Ranks Sum test to examine whether there was a significant difference in RTs between real word trials and pseudoword trials in the LD task. There was a significant difference between mean RTs for words and for pseudowords ($W = 815840$, $p < 0.001$). The descriptive statistics for these differences are presented in Table 41.

Table 41 – Mean RTs and SDs for words and pseudowords in the LD task in Study 2

Trial type	Mean RTs	SDs	Ranges
Word	1150.033	492.382	416-2498
Pseudoword	1383.019	519.751	436-2499

Source: The Author (2024)

I created new subsets of the data for RT analyses with only answers for words and only correct answers for the LD task, for accuracy analyses with only answers for words, for RT analyses on the 5th word (the target one) of the sentences, and for RT analyses on the 6th word of the sentences. I set contrasts and centered all independent variables: cognate status, word position, HC vocabulary, and RH in BP (Brehm; Alday, 2022).

Now I will present the planned inferential statistics tests to confirm or to disconfirm the Hypotheses in this study. Model 1 tested Hypotheses 1 and 4 in relation to RTs in the LD task. Model 2 tested Hypothesis 2 and 5 in relation to accuracy rates in the LD task. Models 3 to 6 tested Hypothesis 3 and 6; model 3 tested them in relation to the 5th word in the sentence in the WSC task; model 4, the 6th word in the sentence in the WSC task; model 7, the 5th word in the sentence in the SSC task; and model 8, the 6th word in the sentence in the SSC task.

Model 1 had the following code: `m21 <- lmer(data = s2LD_wordcorrect, log(rt_clean) ~ status + RHpcent_s + status*RHpcent_s + (1 | participant) + (1 | stimulus))`. Table 42 presents estimates, standard errors, and p-values.

Table 42 – Model 1 of planned analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.018e+00	9.261e-02	1.190e+01	75.781	<2e-16 ***
Status [cognate]	-3.040e-02	1.777e-02	1.113e+02	-1.711	0.0899
RH (%)	-5.040e-02	9.108e-02	1.078e+01	-0.553	0.5913
Status [cognate]:RH	2.297e-03	7.515e-03	1.013e+03	0.306	0.7599
Random effects					
	Variance		SD		
Stimulus	0.03044		0.1745		
Participant	0.10400		0.3225		
Residual	0.06345		0.2519		
N observations	1143				
N stimulus	120				
N participants	13				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

There was no significant main effect or interaction in Model 1. Cognate words across HC and BP did not seem to influence RTs for adult speakers of HC in the BP LD task. Reading habits in BP also did not influence RTs in the task. Thus, Hypotheses 1 and 3 were not confirmed.

Model 2 had the following code: `m22 <- glmer(data = s2LD_word, acc ~ status + RHpcent_s + status*RHpcent_s + (1 | participant) + (1 | stimulus), family = "binomial")`. Table 43 presents estimates, standard errors, and p-values.

Table 43 – Model 2 of planned analyses in Study 2

Fixed effects					
	Estimates	SE	z value		p
(Intercept)	4.01521	0.68646	5.849		4.94e-09 ***
Status [cognate]	0.09969	0.26966	0.370		0.712
RH (%)	0.91750	0.62392	1.471		0.141
Status [cognate]:RH	-0.19214	0.15822	-1.214		0.225
Random effects					
	Variance		SD		
Stimulus	4.168		2.042		
Participant	3.897		1.974		
N observations	1227				
N stimulus	120				
N participants	13				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Similarly, there was no significant main effect or interaction in Model 2. Cognate words across HC and BP did not seem to influence accuracy rates for adult speakers of HC in the BP LD task. Reading habits in BP also did not influence accuracy rates in the task. Thus, Hypotheses 2 and 4 were not confirmed.

Model 3 had the following code: `m23 <- lmer(data = word5W, log(rt_clean) ~ status + RHpcent_s + status*RHpcent_s + (1 | participant) + (1 | word))`. Table 44 presents estimates, standard errors, and p-values.

Table 44 – Model 3 of planned analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.886e+00	9.919e-02	3.105e+01	69.424	<2e-16 ***
Status [cognate]	-5.945e-04	1.647e-02	7.228e+01	-0.036	0.971
RH (%)	2.739e-03	9.680e-02	2.975e+01	0.028	0.978
Status [cognate]:RH	1.836e-02	7.885e-03	1.620e+03	2.328	0.020 *
Random effects					
	Variance		SD		
Word	0.01561		0.1249		
Participant	0.30471		0.5520		
Residual	0.10649		0.3263		
N observations	1723				
N words	75				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 3 presented a significant interaction between cognate status and RH ($p = 0.040$, adjusted). This interaction was further detailed with two other models with only RH as fixed effect: one model with RTs for only cognate words, and the other with RTs for only noncognates. However, in both cases, RH did not reach significance (ps > 0.1). Cognate words across HC and BP did not seem to influence RTs for adult speakers of HC on the 5th word of the WSC task. Reading habits in BP also did not influence RTs in the task. Thus, Hypotheses 3 and 6 were not confirmed.

Model 4 had the following code: `m24 <- lmer(data = word6W, log(rt_clean) ~ previous + RHpcent_s + previous*RHpcent_s + (1`

`| participant) + (1 | word))`. Table 45 presents estimates, standard errors, and p-values.

Table 45 – Model 4 of planned analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.714e+00	9.529e-02	3.358e+01	70.457	<2e-16 ***
Previous [cognate]	-2.570e-04	8.189e-03	1.654e+03	-0.031	0.975
RH (%)	-1.707e-03	8.757e-02	3.003e+01	-0.019	0.985
Previous [cognate]:RH	6.815e-03	6.637e-03	2.070e+03	1.027	0.305
Random effects					
	Variance		SD		
Word	0.01879		0.1371		
Participant	0.27123		0.5208		
Residual	0.09392		0.3065		
N observations	2142				
N words	38				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

There was no significant main effect or interaction in Model 4. Cognate words across HC and BP did not seem to influence RTs for adult speakers of HC on the 6th word of the WSC task. Reading habits in BP also did not influence RTs in the task. Thus, Hypotheses 3 and 6 were not confirmed.

Model 5 had the following code: `m25 <- lmer(data = word5S, log(rt_clean) ~ status + RHpcent_s + status*RHpcent_s + (1 |`

`participant) + (1 | word))`. Table 46 presents estimates, standard errors, and p-values.

Table 46 – Model 5 of planned analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p
(Intercept)	6.23737	0.03148	27.16873	198.164	<2e-16 ***
Status [cognate]	-0.03436	0.01610	74.82823	-2.134	0.0361 *
RH (%)	-0.02560	0.02809	27.70003	-0.912	0.3699
Status [cognate]:RH	0.01869	0.01452	760.96304	1.287	0.1983
Random effects					
	Variance		SD		
Word	0.003615		0.06013		
Participant	0.020208		0.14215		
Residual	0.165559		0.40689		
N observations	812				
N words	80				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 5 presents a significant main effect of cognate status (p = 0.0361). However, this main effect lost significance after adjusting p-values (p = 0.0722). Thus, there was no significant main effect or interaction in Model 5. Cognate words across HC and BP did not seem to influence RTs for adult speakers of HC on the 5th word of the SSC task. Reading habits in BP also did not influence RTs in the task. Thus, Hypotheses 3 and 6 were not confirmed.

Model 6 had the following code: `m26 <- lmer(data = word6S, log(rt_clean) ~ previous + RHpcent_s + previous*RHpcent_s + (1`

`| participant) + (1 | word))`. Table 47 presents estimates, standard errors, and p-values.

Table 47 – Model 6 of planned analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.230e+00	3.358e-02	2.282e+01	185.533	<2e-16 ***
Previous [cognate]	-1.837e-03	1.093e-02	1.043e+02	-0.168	0.867
RH (%)	-1.691e-02	3.097e-02	3.069e+01	-0.546	0.589
Previous [cognate]:RH	-4.158e-03	1.089e-02	1.228e+03	-0.382	0.703
Random effects					
	Variance		SD		
Word	0.0001162		0.01078		
Participant	0.0302897		0.17404		
Residual	0.1484485		0.38529		
N observations	1274				
N words	35				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Similarly, there was no significant main effect or interaction in Model 6. Cognate words across HC and BP did not seem to influence RTs for adult speakers of HC on the 6th word of the SSC task. Reading habits in BP also did not influence RTs in the task. Thus, Hypotheses 3 and 6 were not confirmed.

In sum, none of the Hypotheses in this study was confirmed. This indicates that, for this population sample of adult speakers of HC who migrated to Brazil, cognate words and reading habits in BP do not seem to influence lexical access of words in BP. There seems to be no impact either on words in isolation or on words in sentence context. Also there seems to be no effect difference on words which were spoken or written.

Descriptive statistics after data trimming are the following. Table 48 shows mean RTs, mean accuracy rates and SDs by cognate status in the LD task. Table 49 shows mean RTs and SDs for each word in the sentences in both SC tasks. Table 50 shows mean RTs and SDs for the target word in both SC tasks. Table 51 shows mean RTs and SDs for each RH percentage reported in both SC tasks. Figure 53 shows box and whiskers plots with mean RTs by cognate status in the LD task. Figures 54 and 55 show mean RTs for every word in the SC tasks.

Table 48 – Mean RTs, mean accuracy rates and SDs in the LD task in Study 2

Cognate status	Mean RTs	SDs	Ranges	
Noncognate	1145.386	486.129	494-2498	
Cognate	1101.604	471.017	416-2458	
Cognate status	**Mean accuracy rates (%)**	**SDs**	**Mean accuracy scores**	**Ranges**
Noncognate	0.925	0.263	43.000	0-57
Cognate	0.937	0.242	44.923	1-59

Source: The Author (2024)

Table 49 – Mean RTs and SDs for each word in the SC task in Study 2

Word position	Mean RTs in the SSC task	SDs	Ranges	Mean RTs in the WSC task	SDs	Ranges
1	651.183	364.995	301-2472	900.682	512.244	301-2494
2	601.029	342.153	301-2471	987.019	563.866	302-2485
3	564.055	293.684	301-2316	953.623	574.477	301-2493
4	569.949	290.492	301-2475	916.587	556.652	301-2498
5	577.664	321.723	301-2240	1044.017	594.304	301-2491
6	580.517	311.746	301-2305	903.564	557.543	302-2490
7	645.031	366.381	301-2387	970.721	586.039	301-2486
8	750.925	439.401	301-2460	974.714	587.148	303-2485
9	831.391	440.842	301-2249	1006.735	597.286	306-2485

Source: The Author (2024)

Table 50 – Mean RTs and SDs for the target word in the SC tasks in Study 2

Cognate status of 5th word	Mean RTs in the SSC task	SDs	Ranges	Mean RTs in the WSC task	SDs	Ranges
Noncognate	597.316	335.709	301-2240	1035.091	597.538	303-2490
Cognate	557.121	305.495	301-2151	1059.580	593.901	301-2491

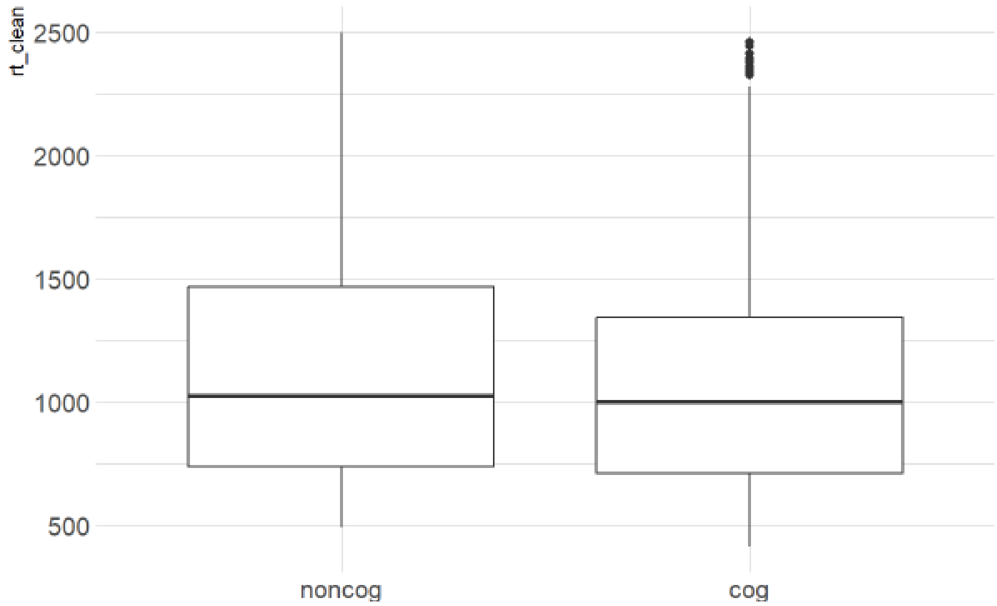
Source: The Author (2024)

Table 51 – Mean RTs and SDs for each RH percentage reported in the SC tasks in Study 2

RH (%)	Mean RTs in the SSC task	SDs	Ranges	Mean RTs in the WSC task	SDs	Ranges
23	749.254	465.368	302-2460	702.432	394.237	301-2437
31	690.977	442.712	301-2461	807.767	504.069	303-2485
37	714.859	362.652	303-2162	1226.575	511.077	324-2485
46	448.881	130.589	301-890	436.752	140.698	301-1081
49	469.474	154.774	302-1057	483.130	260.603	302-2394
51	372.335	62.528	301-607	543.926	110.355	322-1503
54	641.456	322.710	303-2432	992.411	444.508	302-2429
57	618.972	305.103	301-1717	2246.860	246.749	727-2485
60	624.244	321.938	301-2052	1616.497	408.939	588-2493
66	791.152	490.758	302-2471	803.703	308.923	389-2454
69	559.679	258.847	301-2393	840.963	513.095	302-2485
71	672.629	352.419	301-2223	1427.799	584.672	303-2486
74	649.653	413.800	301-2330	1182.713	503.088	388-2493
77	580.247	336.294	301-2196	696.990	388.235	303-2486
80	627.059	351.675	301-2475	886.549	519.176	302-2498
83	463.275	163.281	301-1010	604.083	223.431	309-1781
86	649.626	381.858	302-1976	452.266	290.406	301-2491
89	573.982	317.140	302-2366	886.503	449.088	319-2470
97	776.561	539.565	303-2443	935.114	479.805	321-2485

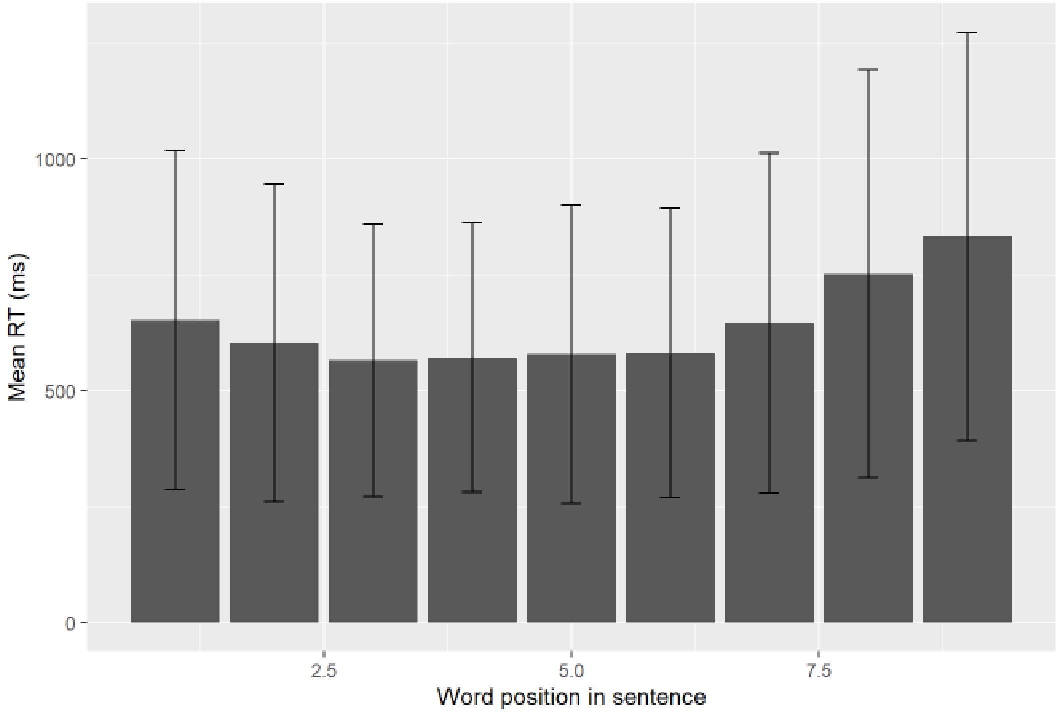
Source: The Author (2024)

Figure 53 – Boxplot of RTs by cognate status in the LD task in Study 2



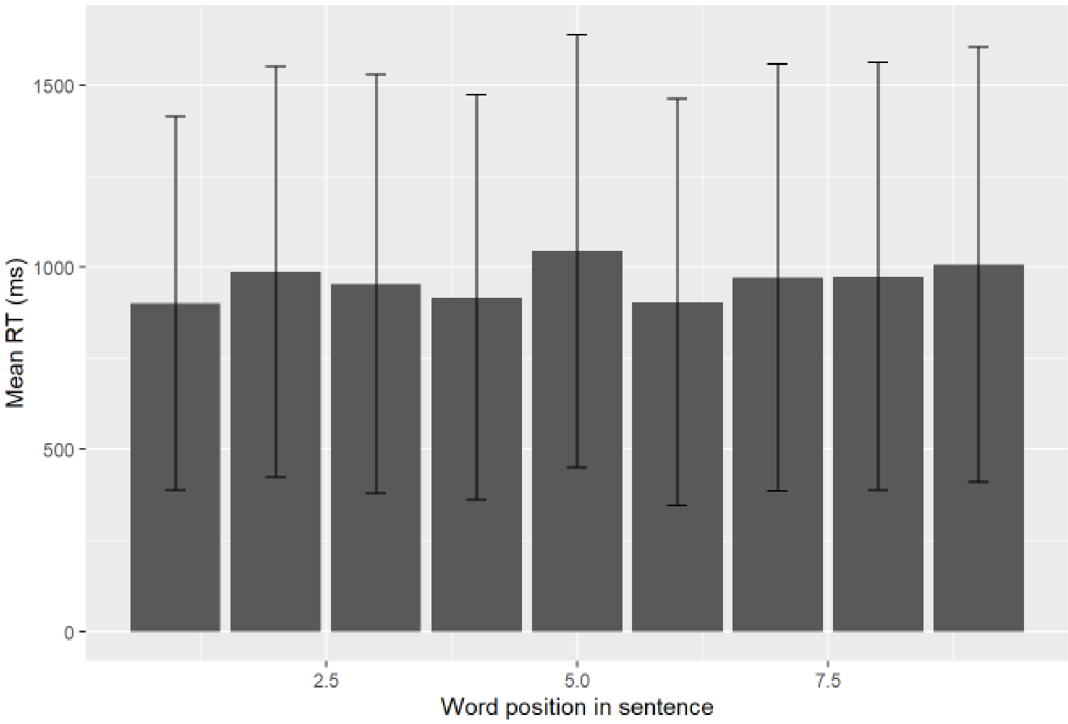
Source: The Author (2024)

Figure 54 – Mean RTs for every word in the SSC task in Study 2



Source: The Author (2024)

Figure 55 – Mean RTs for every word in the WSC task in Study 2



Source: The Author (2024)

Now I will present exploratory analyses. These include further analyses to make sure the absence of a cognate effect in planned analysis does not stem from other confounding variables. Exploratory analyses also include tests ran on residual RTs after trimming the auditory stimuli duration from RTs in the SSC task (van Dijk; Dijkstra; Unsworth, 2022). This will be detailed later.

When selecting stimuli for the LD and SC tasks, it was not possible to control for word length. I performed further analyses to examine whether length could have interacted with cognate status and lead to the absence of cognate effect. Model 7 had cognate status and word length as fixed effects and tested their effects on RTs in the LD task. Model 8 tested the same fixed effects on RTs for the target word in the WSC task. Model 9 tested the same fixed effects on RTs for the target word in the SSC task.

Model 7 presented a significant main effect of length (p = 7.46^{e-14}) and a marginal interaction between cognate status and length (p = 0.0561). The marginal interaction was discarded since the p-value would increase after adjustments. Table 52 presents estimates, standard errors, and p-values.

Table 52 – Model 7 of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p
(Intercept)	6.617740	0.101791	21.005235	65.013	<2e-16 ***
Status [cognate]	0.060953	0.051282	109.305878	1.189	0.2372
Letters	0.063607	0.007444	111.755964	8.544	7.46e-14 ***
Status [cognate]:Letters	-0.014355	0.007437	111.319700	-1.930	0.0561
Random effects					
	Variance		SD		
Stimulus	0.01573		0.1254		
Participant	0.09982		0.3160		
Residual	0.06607		0.2570		
N observations	1227				
N stimulus	120				

N participants	13				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 7 shows that participants took longer to respond to words in the LD task when they were longer. Word length is known to influence word processing: longer words take longer to process than shorter words (Cortese; Balota, 2012; Yap; Balota, 2015). This effect may be more intense for readers who have less efficient decoding skills (Barton *et al*., 2014; Gerth; Festman, 2021). Word length could have obscured the cognate effect if only shorter words had shown differences in RTs between cognate and noncognate ones. Also, it is not possible to state that cognate status facilitated or inhibited the processing of longer words. Thus, it seems improbable that word length may have obscured the cognate effect. Table 53 presents mean RTs for each word length and each cognate status.

Table 53 – Mean RTs for each word length and each cognate status in the LD task in Study 2

Cognate status	Word length	Mean RTs	SDs	Ranges
Noncognate	3	880.435	301.987	528-1690
Cognate	3	792.455	232.702	533-1386
Noncognate	4	886.547	287.017	494-1546
Cognate	4	953.938	363.423	480-2010
Noncognate	5	1056.040	442.148	514-2426
Cognate	5	1055.194	456.587	511-2442
Noncognate	6	1143.594	500.877	511-2472
Cognate	6	1055.056	444.130	507-2337
Noncognate	7	1167.532	498.157	519-2498
Cognate	7	1119.528	480.469	430-2458
Noncognate	8	1293.442	525.614	555-2394
Cognate	8	1172.009	479.743	416-2393
Noncognate	9	1293.841	508.276	508-2360
Cognate	9	1060.207	414.709	505-1930
Noncognate	10	1402.000	469.315	641-2379
Cognate	10	1463.889	523.136	592-2171
Noncognate	11	1677.571	516.903	984-2324
Cognate	11	1142.800	583.667	559-2348
Noncognate	12	1467.600	404.301	1040-2106
Cognate	12	1619.636	561.663	795-2411

Source: The Author (2024)

Model 8 also presented a significant main effect of length for the target word in the WSC task (p = 1.7e-12), which may be observed in Table 54. This indicates that participants took longer to respond to the target words in the WSC task when they were longer. However, the same was not true for the SSC task.

Table 54 – Model 8 of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p
(Intercept)	6.519319	0.105921	43.277145	61.549	<2e-16 ***
Status [cognate]	-0.023293	0.044117	65.918095	-0.528	0.599
Length	0.054605	0.006366	69.313536	8.578	1.7e-12 ***
Status [cognate]:Length	0.002080	0.006335	67.976710	0.328	0.744
Random effects					
	Variance		SD		
Word	0.005358		0.0732		
Participant	0.296588		0.5446		
Residual	0.106774		0.3268		
N observations	1723				
N words	75				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 9 showed no significant main effects or interactions (Table 55). This indicates that, in the SSC task, neither cognate status nor word length had influences over the time participants took to respond to spoken target words. This may also be observed in Tables 56 and 57, which present mean RTs for each word length and each cognate status in both SC tasks.

Table 55 – Model 9 of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.2194290	0.0684807	89.1893827	90.820	<2e-16 ***
Status [cognate]	-0.0375570	0.0625525	72.5475156	-0.600	0.550
Length	0.0024047	0.0095635	78.2010921	0.251	0.802
Status [cognate]:Length	0.0004624	0.0095317	76.9975678	0.049	0.961
Random effects					
	Variance		SD		
Word	0.004326		0.06578		
Participant	0.020009		0.14145		
Residual	0.165561		0.40689		
N observations	812				
N words	80				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Table 56 – Mean RTs for each word length and each cognate status in the SSC task in Study 2

Cognate status	Word length	Mean RTs	SDs	Ranges
Noncognate	4	621.238	390.000	301-2366
Cognate	4	550.696	360.741	301-2151
Noncognate	5	557.901	226.811	301-1184
Cognate	5	595.791	254.042	315-1284
Noncognate	6	568.328	314.707	307-1960
Cognate	6	489.524	208.188	306-1442
Noncognate	7	662.723	378.911	301-1796
Cognate	7	564.198	301.868	301-1572
Noncognate	8	538.093	318.050	301-2240
Cognate	8	629.029	424.288	301-2013
Noncognate	9	670.632	397.745	307-1906
Cognate	9	587.125	319.237	322-1182

Noncognate	10	551.600	163.906	390-768
Cognate	10	543.235	342.804	302-1788
Cognate	11	599.546	269.841	322-1133

Source: The Author (2024)

Table 57 – Mean RTs for each word length and each cognate status in the WSC task in Study 2

Cognate status	Word length	Mean RTs	SDs	Ranges
Noncognate	2	827.280	579.213	312-2293
Noncognate	4	973.665	590.465	307-2462
Cognate	4	934.654	580.311	312-2453
Noncognate	5	995.395	597.212	308-2485
Cognate	5	1020.235	606.750	301-2485
Noncognate	6	1013.438	584.844	303-2468
Cognate	6	992.590	561.954	305-2485
Noncognate	7	1074.520	614.764	304-2490
Cognate	7	1092.196	583.020	304-2453
Noncognate	8	1097.082	626.726	318-2485
Cognate	8	1131.109	603.498	303-2491
Noncognate	9	1055.815	539.938	312-2372
Cognate	9	1149.540	620.489	309-2394
Noncognate	10	1204.409	530.799	484-2484
Cognate	10	1010.289	577.786	329-2468
Noncognate	11	1182.651	600.267	322-2485

Source: The Author (2024)

Therefore, adult speakers of HC took longer to respond to longer BP words during visual recognition both in isolation and in sentence context. Despite this effect, cognate words do not seem to have affected the processing of longer words, and word length does not appear to have influenced the lack of a cognate effect.

Another exploratory analysis I investigated involved reading habits and the processing of all words in the sentences. The interaction between RH and cognate status in Model 3 appears to be spurious. However, since planned models examined only target words in the SC tasks and did not consider possible RH effects to entire sentences, I included two models to test whether RH would influence RTs for all words in the sentences. Table 58 presents Model 10, which investigated RTs in the WSC task, and Table 59 presents Model 11, which investigated RTs in the SSC task.

Table 58 – Model 10 of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value

(Intercept)	6.821545	0.088209	30.408285	77.33	<2e-16 ***
RH (%)	0.004148	0.083591	29.996915	0.05	0.961
Random effects					
	Variance		SD		
Word	0.01816		0.1348		
Participant	0.24607		0.4961		
Residual	0.12153		0.3486		
N observations	15461				
N words	363				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Table 59 – Model 11 of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	6.28465	0.02950	33.93663	213.064	<2e-16 ***
RH (%)	-0.01360	0.02655	29.51616	-0.512	0.612
Random effects					
	Variance		SD		
Word	0.01617		0.1272		
Participant	0.02458		0.1568		
Residual	0.17052		0.4129		
N observations	8667				
N words	363				

N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

However, in both models, RH did not reach significance. This indicates that, for this population sample of adult speakers of HC who migrated to Brazil, RH in BP does not seem to influence RTs when reading or listening to words in sentences in BP.

Now I will describe further exploratory analyses which were inspired by the ones performed by van Dijk, Dijkstra, and Unsworth (2022). They used a self-paced sentence listening task to investigate syntactic cross-language influences. Since the segments participants listened to had different durations, they followed the procedures recommended by Marinis (2010). I performed further exploratory analyses using this procedure on the SSC task because word length of stimuli varied.

First, residual RTs were computed. For this, the duration of each audio file was removed from its respective RT for all participants so that the residual RT would indicate the time participants took to respond after the end of the audio. This means that if a participant pressed the button before the end of the audio, the residual RT would be a negative number. In order to solve this, a constant was added to all RT values. This constant was the lowest RT value transformed to a positive number plus one. The lowest RT value for each participant was calculated; it was -1,522 ms. This number was transformed into a positive number (1,522) and was added one (1,523). This constant was aggregated to all residual RTs in both task modalities.

After calculating the residual RTs in the SSC and adding the constant to both tasks, data trimming was carried out similarly to the steps reported in the planned analyses above. The only descriptive statistic which changed from planned analyses to this residual RTs one is RT values. Table 60 presents mean residual RTs, SDs, and ranges for both SC tasks.

Table 60 – Mean residual RTs, SDs, and ranges for both SC tasks in Study 2

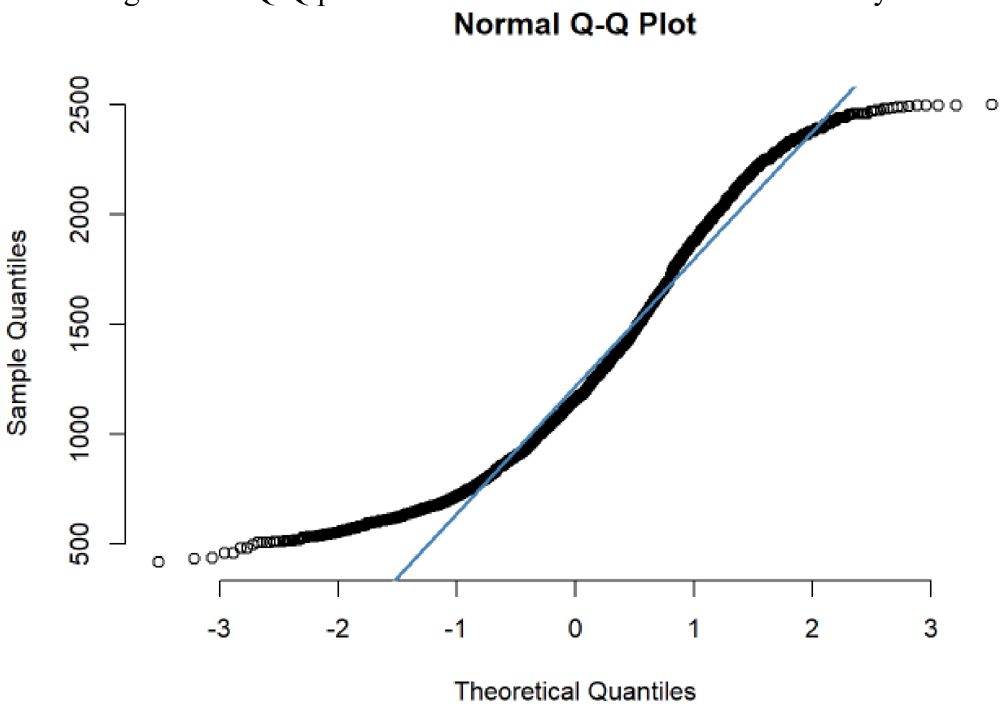
Task modality	Mean RTs (ms)	SDs	Range (ms)
Spoken Sentence Comprehension task	2151.448	14613.160	1—1828878
Written Sentence Comprehension task	7222.174	566842.900	1524—84235188

Source: The Author (2024)

Data from participants who had accuracy rates lower than 70% in the LD task was removed from the LD task analyses. Also, data from participants who had less than 50% of accuracy in the comprehension questions in the WSC task was removed from both SC tasks analyses. RT values deviating 2.5 SDs below or above the mean for each participant were removed. RT values shorter than 1,823 ms (300 ms + 1,523) and longer than 4,023 ms (2,500 ms + 1523) were removed from the SC tasks files. RT values shorter than 300 ms and longer than 2,500 ms were removed from the LD task file. Data was also removed according to scores in the HC vocabulary test. These steps removed the same percentage of the data as in the planned analyses.

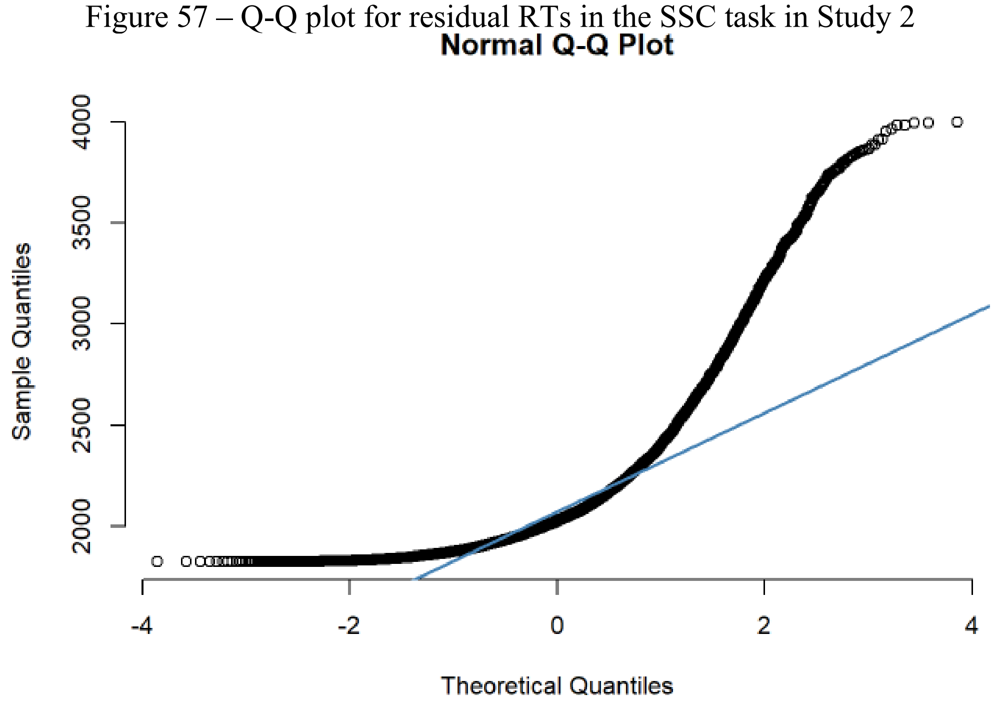
The distribution of RTs was also not normal (LD task: $A = 40.114$, $p < 0.001$; SSC task: $A = 561.4$, $p < 0.001$; WSC task: $A = 636.11$, $p < 0.001$), which can be verified with Figures 56 to 61 which show Q-Q plots and histogram plots.

Figure 56 – Q-Q plot for residual RTs in the LD task in Study 2



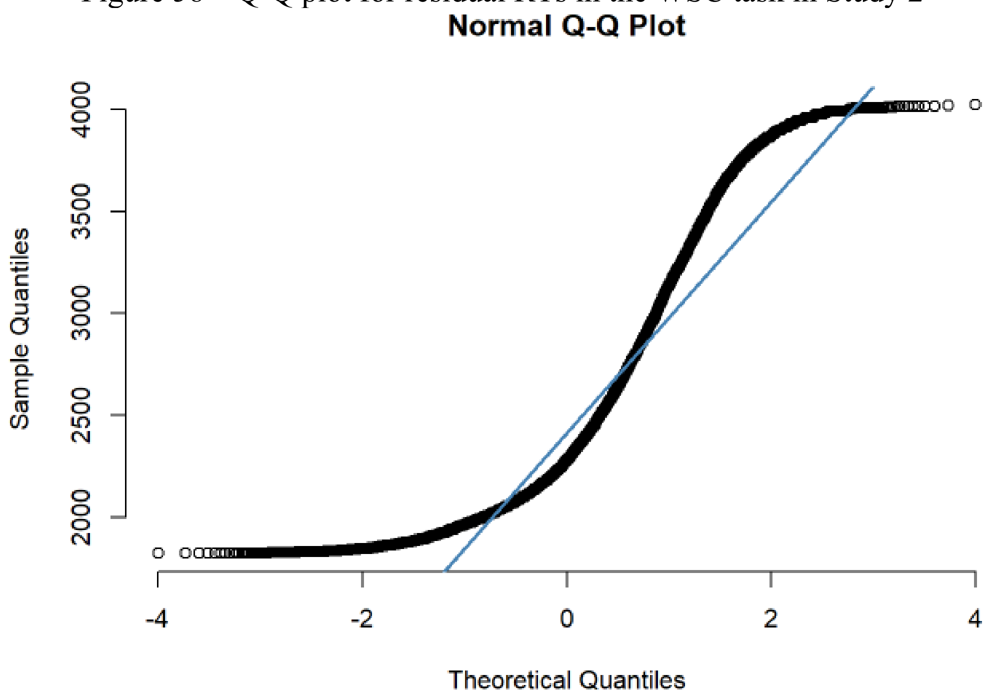
Source: The Author (2024)

Figure 57 – Q-Q plot for residual RTs in the SSC task in Study 2



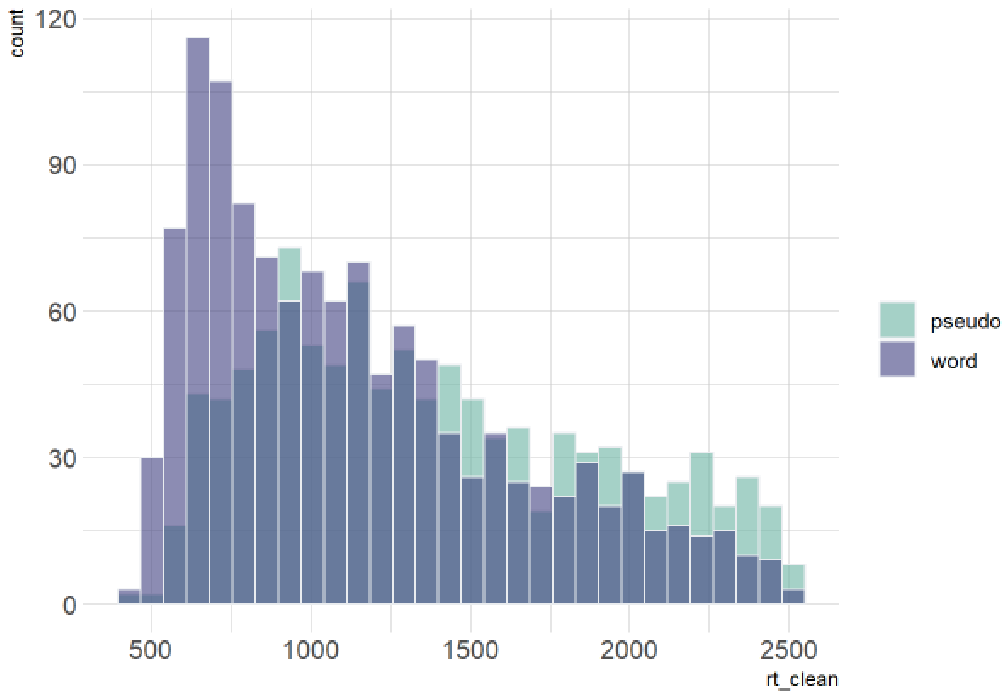
Source: The Author (2024)

Figure 58 – Q-Q plot for residual RTs in the WSC task in Study 2



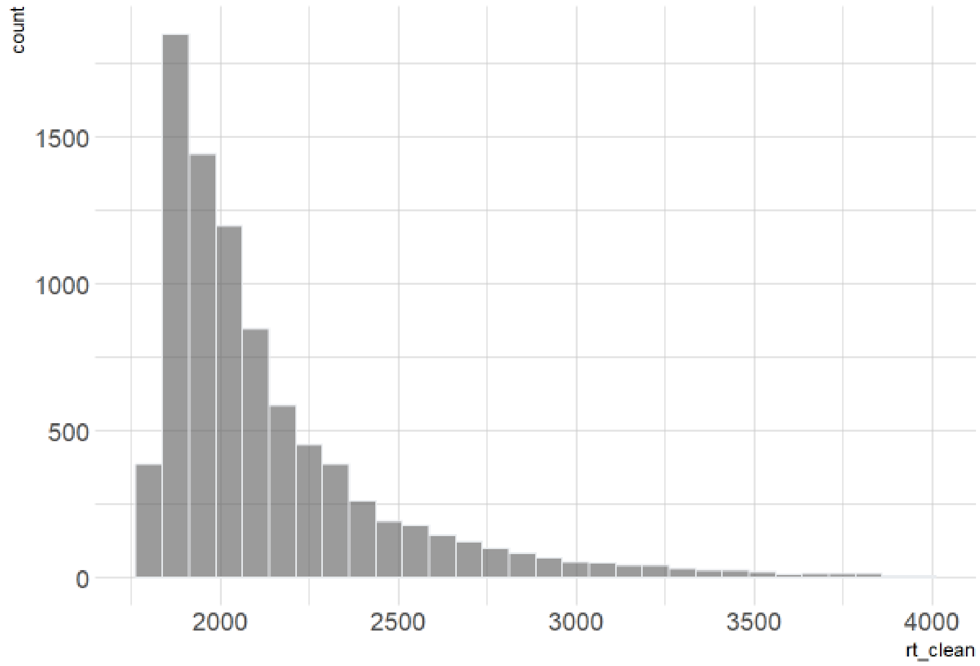
Source: The Author (2024)

Figure 59 – Histogram plot for residual RTs in the LD task in Study 2



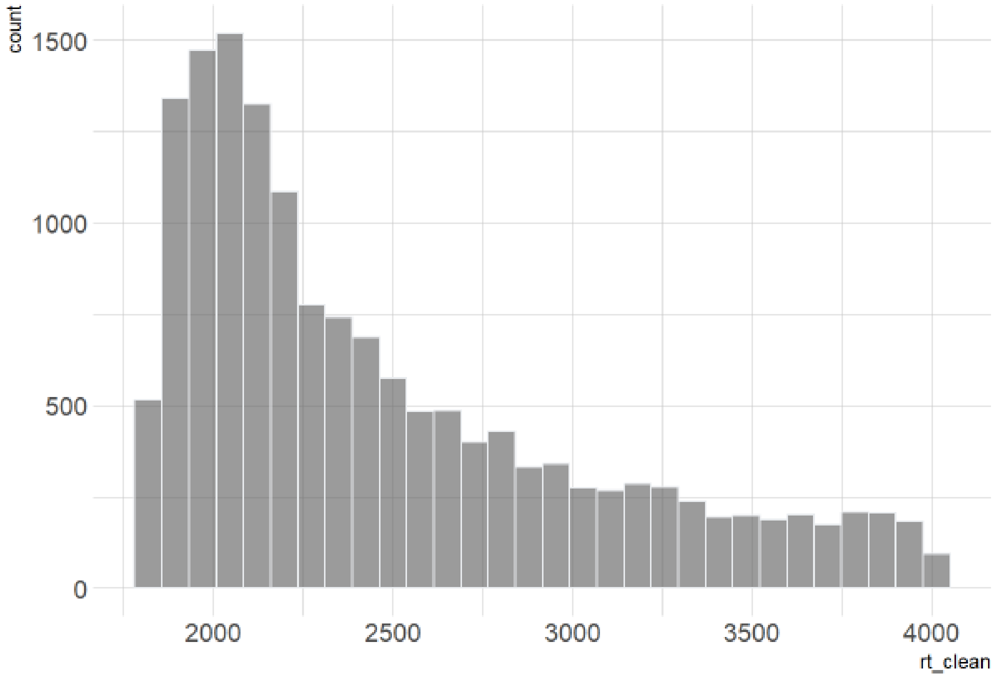
Source: The Author (2024)

Figure 60 – Histogram plot for residual RTs in the SSC task in Study 2



Source: The Author (2024)

Figure 61 – Histogram plot for residual RTs in the WSC task in Study 2



Source: The Author (2024)

Thus, I used a Wilcoxon Ranks Sum test to examine whether there was a significant difference in RTs between real word trials and pseudoword trials in the LD task. Indeed, RTs to real words were shorter than to pseudowords ($W = 815840$, $p < 0.001$). Descriptive statistics may be examined in Table 61.

Table 61 – Mean RTs and SDs for words and pseudowords in the LD task in Study 2

Trial type	Mean RTs	SDs	Ranges
Word	1150.033	492.382	416-2498
Pseudoword	1383.019	519.752	436-2499

Source: The Author (2024)

Before removing data from participants who had accuracy scores lower than 70% in the LD task, the mean age was 32.14. After that, mean age for the LD data was 29.92. Before removing data from participants who had accuracy scores lower than 50% in the SC tasks, the mean age was 32.14. After that, mean age for the SC data was 32.09. Median age (33) was the same before and after accuracy trimming for all tasks.

Contrasts were set and variables were centered following the steps in the planned analyses. Five mixed-effect linear models were used to test Hypotheses about RTs in relation to residual RTs. They were constructed exactly like in the planned analyses described above.

Model 1 had the following code: `m21r <- lmer(data = s2LD_wordcorrect, log(rt_clean) ~ status + RHpcent_s + status*RHpcent_s + (1 | participant) + (1 | stimulus))`. Table 62 presents estimates, standard errors, and p-values. Results show that there was no significant main effect or interaction in the model for RTs in the LD task.

Table 62 – Model 1 on residual RTs of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.018e+00	9.261e-02	1.190e+01	75.781	<2e-16 ***
Status [cognate]	-3.040e-02	1.777e-02	1.113e+02	-1.711	0.0899
RH (%)	-5.040e-02	9.108e-02	1.078e+01	-0.553	0.5913
Status [cognate]:RH	2.297e-03	7.515e-03	1.013e+03	0.306	0.7599
Random effects					
	Variance		SD		
Stimulus	0.03044		0.1745		
Participant	0.10400		0.3225		
Residual	0.06345		0.2519		
N observations	1143				
N stimulus	120				
N participants	13				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 2 had the following code: `m22r <- lmer(data = word5W, log(rt_clean) ~ status + RHpcent_s + status*RHpcent_s + (1 | participant) + (1 | word))`. Table 63 presents estimates, standard errors, and p-values. Results show a significant interaction between cognate status and RH for the target word in the WSC task, just like in Model 3 in planned analyses.

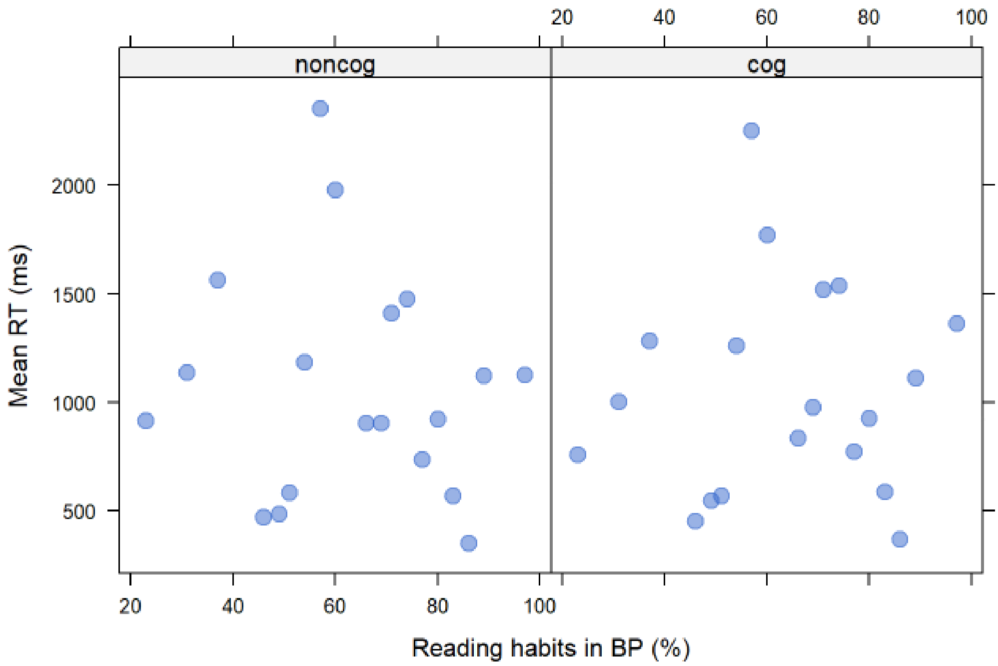
Table 63 – Model 1 on residual RTs of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.869e+00	3.825e-02	3.102e+01	205.706	<2e-16 ***
Status [cognate]	-8.583e-04	6.273e-03	7.239e+01	-0.137	0.8915
RH (%)	-1.115e-03	3.734e-02	2.975e+01	-0.030	0.9764
Status [cognate]:RH	8.664e-03	3.023e-03	1.620e+03	2.867	0.0042 **
Random effects					
	Variance		SD		
Word	0.002255		0.04749		
Participant	0.045354		0.21297		
Residual	0.015649		0.12510		
N observations	1723				
N stimulus	75				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

This interaction was further analyzed using two models, one with RTs only for cognate words and the other with RTs only for noncognates. However, in both models, RH did not reach significance ($ps > 0.7$). This is probably a spurious result. Figure 62 demonstrates how RT values may be distributed according to their respective RH percentage and cognate status.

Figure 62 – Mean RTs by cognate status and RH (%) for the target word in the WSC task in Study 2



Source: The Author (2024)

Model 3 had the following code: `m23r <- lmer(data = word6W, log(rt_clean) ~ previous + RHpcent_s + previous*RHpcent_s + (1 | participant) + (1 | word))`. Table 64 presents estimates, standard errors, and p-values. Results show that there was no significant main effect or interaction in the model for RTs for the 6th word in the sentences in the WSC task.

Table 64 – Model 3 on residual RTs of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.803e+00	3.586e-02	3.377e+01	217.594	<2e-16 ***
Previous [cognate]	1.154e-03	3.063e-03	1.702e+03	0.377	0.706
RH (%)	-4.371e-03	3.290e-02	3.003e+01	-0.133	0.895
Previous [cognate]:RH	3.037e-03	2.477e-03	2.070e+03	1.226	0.220
Random effects					
	Variance		SD		

Word	0.002802		0.05293		
Participant	0.038301		0.19571		
Residual	0.013077		0.11436		
N observations	2142				
N words	38				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 4 had the following code: `m24r <- lmer(data = word5S, log(rt_clean) ~ status + RHpcent_s + status*RHpcent_s + (1 | participant) + (1 | word))`. Table 65 presents estimates, standard errors, and p-values. Results show that there was no significant main effect or interaction in the model for RTs for the target word in the sentences in the SSC task. There was a marginal main effect of cognate status (p = 0.056). This p-value was not considered because it would become even less significant after adjustments.

Table 65 – Model 4 on residual RTs of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.637771	0.009382	27.305186	814.046	<2e-16 ***
Status [cognate]	-0.009674	0.005060	75.058025	-1.912	0.0597
RH (%)	-0.007948	0.008358	27.807541	-0.951	0.3498
Status [cognate]:RH	0.006576	0.004540	762.277415	1.449	0.1479
Random effects					
	Variance		SD		
Word	0.0003749		0.01936		

Participant	0.0017090		0.04134		
Residual	0.0162015		0.12729		
N observations	812				
N words	80				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

Model 5 had the following code: `m25r <- lmer(data = word6S, log(rt_clean) ~ previous + RHpcent_s + previous*RHpcent_s + (1 | participant) + (1 | word))`. Table 66 presents estimates, standard errors, and p-values. Results show that there was no significant main effect or interaction in the model for RTs for the 6th word in the sentences in the SSC task.

Table 66 – Model 5 on residual RTs of exploratory analyses in Study 2

Fixed effects					
	Estimates	SE	df	t	p value
(Intercept)	7.634e+00	9.883e-03	3.011e+01	772.452	<2e-16 ***
Previous [cognate]	-4.248e-04	3.398e-03	1.247e+03	-0.125	0.901
RH (%)	-5.456e-03	9.164e-03	3.126e+01	-0.595	0.556
Previous [cognate]:RH	-2.186e-03	3.403e-03	1.248e+03	-0.642	0.521
Random effects					
	Variance		SD		
Word	0.000000		0.0000		
Participant	0.002601		0.0510		
Residual	0.014491		0.1204		
N observations	1274				

N words	35				
N participants	32				
Significance codes	0 = ***	0.001 = **	0.01 = *	0.05 = .	0.1 =

Source: The Author (2024)

When comparing the results from Models 1 to 6 in the planned analyses and from Models 1 to 5 in the residual RTs exploratory analyses, no difference in main effects or interactions is found. In sum, exploratory analyses with residual RTs did not add any different results from the ones presented in planned analyses. Finally, in this study, it seems that adult speakers of HC who migrated to Brazil and use BP in their everyday lives do not demonstrate effects from cognate words across HC and BP and from RH in BP when they are reading or listening to isolated words or to words in sentence context in BP.

In the next section, I will discuss these results in relation to the literature and to potential alternative perspectives.

4.3 DISCUSSION

In this section, I will refer to the research questions, objectives and hypotheses and sum up the results before discussing them in light of the literature.

Study 2 investigated whether adults who are native speakers of HC show effects of cross-linguistic interaction between HC and BP on lexical access during comprehension of spoken and written BP. Study 2 was guided by the following research questions:

1. Do cognate words across HC and BP, in comparison with noncognate words, facilitate visual lexical access in BP?
2. Do cognate words across HC and BP, in comparison with noncognate words, facilitate visual and auditory sentence comprehension in BP?
3. Do reading habits (RH) in BP influence visual lexical access in BP?
4. Do RH in BP influence visual and auditory sentence comprehension in BP?

In order to answer these research questions, Study 2 pursued the following specific objectives:

1. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during a visual lexical decision task in BP;
2. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce higher accuracy rates during a visual lexical decision task in BP;
3. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during visual and auditory sentence comprehension tasks in BP;
4. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during a visual lexical decision task in BP;
5. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce higher accuracy rates during a visual lexical decision task in BP;
6. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during visual and auditory sentence comprehension tasks in BP.

Objectives 1 and 2 are related to research question 1; Objective 3 is related to research question 2; Objectives 4 and 5 are related to research question 3, and Objective 6 is related to research question 4. The Hypotheses for Study 2 are:

1. Cognate words, in comparison with noncognate words, produce shorter RTs during a visual lexical decision task in BP;
2. Cognate words, in comparison with noncognate words, produce higher accuracy rates during a visual lexical decision task in BP;
3. Cognate words, in comparison with noncognate words, produce shorter RTs during both visual and auditory sentence comprehension tasks in BP;
4. More frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during a visual lexical decision task in BP;
5. More frequent RH in BP, in comparison with less frequent RH, produce higher accuracy rates during a visual lexical decision task in BP;
6. More frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during both visual and auditory sentence comprehension tasks in BP.

In order to test these hypotheses, I invited adult speakers of HC who were living in Brazil to perform 4 tasks in BP and one in HC. Data from a total of 35 participants was included in the analyses. These participants completed a visual lexical decision (LD) task with cognate and noncognate words across HC and BP, one visual and one auditory self-paced sentence comprehension (SC) task with cognate and noncognate words across HC and BP, one questionnaire about reading habits in BP, and one HC receptive vocabulary test. The lexical decision task and the sentence comprehension tasks were intended to test whether cognate words across HC and BP would have any significant influences in the lexical access of words in BP recognized visually and auditorily. The LD task would test lexical access of isolated words while the SC tasks would test lexical access of words inserted in sentential context. The reading habits (RH) questionnaire was intended to measure the frequency of reading materials in BP to verify whether lexical access would be modulated by it.

Results from planned analyses showed no significant nor reliable main effects or interactions involving cognate status in any of the three tasks testing the cognate effect. When analyzing RTs in the LD task, there was a marginal effect of cognate status (p = 0.0899), which was not considered further. In the written SC task, a significant interaction between cognate status and RH was observed for RTs on the 5th word — the target word — of sentences (p = 0.0400, adjusted). However, when further analyzing this interaction, there was no effect of RH. Also, the significant main effect of cognate status on RTs in the spoken SC task was not maintained after adjusting p-values (p = 0.0722, adjusted). These results indicate that the presence of cognate words across HC and BP did not significantly facilitate and did not significantly inhibit lexical access of words in BP for adult speakers of HC and BP.

In addition, there were no significant main effects or interactions of RH in BP. The only significant interaction involving RH showed no main effects of RH after further analyses. Thus, this interaction is not reliable. These results indicate that RH of materials in BP do not seem to influence BP word recognition of adult speakers of HC and BP.

In order to further explore the absence of cognate effects and of RH in this study, I performed exploratory analyses on word length and on RH in relation to the entire sentences. The investigation about word length will be presented in Section 3.3.2. Now I will present the analyses on RH and the entire sentences.

In the SC tasks, planned analyses investigated the effect of cognate status and of RH on the 5th and the 6th words of the sentence only (Tables 44 to 47). The 5th word was always the target word, which would be either cognate or noncognate. As explained in Section 3.1.3, all the other words in the sentences were selected from a set of words classified was noncognate across HC and BP. This means that, in planned analyses, the effect of RH was only tested in relation to RTs for part of the sentences, not the entire sentence. So, two models were used to verify whether participants' RH would have any influence on RTs while self-paced reading or self-paced listening to sentences in BP. However, there were no significant effects in either model (*p*s > 0.6). This may indicate that how frequently participants read in BP does not seem to affect the speed of lexical access in BP.

More exploratory analyses were conducted employing the data trimming strategy by van Dijk, Dijkstra, and Unsworth (2022). The goal of this strategy was to calculate residual RTs for the spoken SC task, that is, the time participants took to respond to the auditory stimuli subtracting the duration of each auditory stimuli. After conducting this extra trimming, the same models used in planned analyses were run. However, these exploratory analyses presented no significant differences in relation to the planned ones.

Thus, results showed that, in this study, it seems that adult speakers of HC who migrated to Brazil and use BP in their everyday lives do not demonstrate effects from cognate words across HC and BP when they are reading or listening to isolated words or to word embedded in sentences in BP. Moreover, it seems that visual or auditory lexical access in BP for this sample of participants is not influenced by how frequently they read materials in BP. This means that all 6 Hypotheses in this study were not confirmed.

Hypotheses 1 and 2 stated that cognate words, in comparison with noncognate words, are associated with shorter RTs and higher accuracy rates during a visual LD in BP. However, they were completely disconfirmed. No cognate effect was seen either in RTs or accuracy rates. This result diverges from most of the evidence available in the literature (e.g., Lemhöfer; Spalek; Schriefers, 2008; Mulder; Dijkstra; Baayen, 2015; Mutendam *et al*., 2022; Szubko-Sitarek, 2011; Cassol Rigatti; Arêas da Luz Fontes, 2022).

For example, Lemhöfer, Spalek, and Schriefers (2008) and Szubko-Sitarek (2011) identified shorter RTs for cognate words in relation to noncognates. Mulder *et al.* (2014) and Mulder, Dijkstra, and Baayen (2015) observed shorter RTs for identical cognates. In opposition, Mulder *et al*. (2022) observed that cognate words had no effect on RTs in relation to noncognates. However, Mulder *et al*. (2022) also verified that cognate words produce higher accuracy scores in an auditory LD task in comparison with noncognates, as did Lemhöfer, Spalek, and Schriefers (2008). Similarly, Cassol Rigatti and Arêas da Luz Fontes (2022) observed that high accuracy rates in visual word recognition were linked to cognate words when these words shared a dominant meaning across languages. Contrarily, the study by Comesaña *et al.* (2021) showed that cognate words either decrease accuracy rates or have no effect on word recognition in relation to noncognates. Moreover, I could only find one comprehension study focusing on cross-language activation in adult bilingual speakers of a heritage language. In this study (Muntendam *et al.,* 2022), participants recognized cognate words faster than noncognates in an auditory task.

Thus, although there is evidence that cognate words may cause interference or no effect at all, most studies point to a facilitation effect in relation to noncognate words. Hypothesis 1 diverged from Lemhöfer, Spalek, and Schriefers (2008), Szubko-Sitarek (2011), Mulder *et al.* (2014), Mulder, Dijkstra, and Baayen (2015), and Muntendam *et al.* (2022) and converged with Mulder *et al*. (2022). Hypothesis 2 diverged from Lemhöfer, Spalek, and Schriefers (2008), Mulder *et al*. (2022), Cassol Rigatti and Arêas da Luz Fontes (2022) and converged with Comesaña *et al.* (2021).

Hypothesis 3 stated that cognate words, in comparison with noncognate words, are associated with shorter RTs during visual and auditory SC tasks in BP. However, this hypothesis was disconfirmed. This result diverges from most of the evidence available in the literature (Bultena; Dijkstra; van Hell, 2014; Joss; Virtue, 2010; Schwartz; Kroll, 2007; Soares *et al.* 2019; Toassi; Mota; Teixeira, 2020; van Assche *et al.*, 2009; van Assche *et al.*, 2011).

For example, Toassi, Mota, and Teixeira (2020) observed that cognate words produced shorter eye movements during sentence reading in comparison with noncognates. Bultena, Dijkstra, and van Hell (2014) observed shorter RTs for cognate words in comparison with noncognates during a self-paced sentence reading task. In addition, Chambers and Cooke (2009) reported interference effects from interlingual homophones in a sentence listening task, and van Dijk, Dijkstra, and Unsworth (2022) observed facilitative cross-language syntactic effects in a self-paced listening task. The studies by Chambers and Cooke (2009) and van Dijk, Dijkstra, and Unsworth (2022) point to cross-language activation effects in sentence listening tasks. However, no evidence of language co-activation was seen in this study for any modality of the SC task. Contrarily, Hypothesis 3 agrees with results from Egan *et al.* (2019), Allen, Conklin, and Miwa (2021), and Lijewska (2023) in that they have not seen the cognate effect in sentential context.

Thus, although there is evidence that cognate words may cause interference or no effect at all in sentence context, most studies point to a facilitation effect in relation to noncognate words. Hypothesis 3 diverged from Schwartz and Kroll (2007), van Assche *et al*. (2009), Chambers and Cooke (2009), Joss and Virtue (2010), Van Assche *et al*. (2011); Bultena, Dijkstra, and van Hell (2014), Soares *et al*. (2019), Toassi, Mota, and Teixeira (2020), van Dijk, Dijkstra, and Unsworth (2022) and converged with Egan *et al.* (2019), Allen, Conklin, and Miwa (2021), and Lijewska (2023).

Hypotheses 4 and 5 stated that more frequent RH in BP are associated with shorter RTs and higher accuracy rates during a visual LD in BP. However, there was no reliable effect of RH in either of those measures. These results diverge from most of the studies on RH and reading experience mentioned in Section 2.5 (Hassan *et al.* 2021; Pratheeba; Krashen, 2013; Rudell; Hu, 2010; Santos-Díaz, 2017; Sun; Bornstein; Esposito, 2021).

For example, Rudell and Hu (2010) observed that higher experience with words contributes to higher speed of lexical access. Butler (2011) identified that reading more Japanese books outside of school was associated with better *kanji* reading skills. Pratheeba and Krashen (2013) found a positive correlation between performance in a vocabulary test and reading frequency. Sun, Bornstein, and Esposito (2021) observed that frequency of reading at home was the strongest predictor of word reading skill for young children. Santos-Díaz (2017) found a strong correlation between reading frequency and active and passive vocabulary in foreign language. And Hassan *et al*. (2021) observed that reading habits were positively correlated with reading achievement in a foreign language. Because of evidence such as this, participants who reported reading more frequently in BP were expected to display shorter RTs and higher accuracy rates when recognizing written words in BP.

Similarly, Hypothesis 6 stated that more frequent RH in BP are associated with shorter RTs during visual and auditory SC tasks in BP. However, there was no effect of RH on RTs in any of the modalities of the SC task. This result diverges from most of the studies on RH and reading experience mentioned in Section 2.5 (Hassan *et al.* 2021; Rudell; Hu, 2010; Santos-Díaz, 2017; Sun; Bornstein; Esposito, 2021). The expectation that participants who reported reading more frequently in BP would show shorter RTs when reading and listening to sentences in BP came from the same evidence as Hypotheses 4 and 5 above.

Contrarily, Artieda (2007) observed that only intermediate learners were influenced by reading habits. However, participants in this study showed no reliable influence from RH and presented high error rates which may be verified by the percentage of the data which had to be excluded from inferential analyses (more than 50%). This might indicate that participants in this study were not at an intermediate level of proficiency in BP despite reporting good proficiency levels in the language history questionnaire (Table 36). In addition, Caylak Toplu and Erten (2023) concluded that L1 reading habits predict L1 vocabulary size, which in turn predicts L2 vocabulary size, but that L2 reading habits do not explain L2 vocabulary size. This could mean an indirect relationship between L1 and L2 reading habits and L2 reading and listening skills.

In sum, data from the population sample in this study did not confirm any of the Hypothesis proposed. The absence of any cognate effect — either a facilitatory or an inhibitory one, either in RTs or in accuracy rates, either for isolated words or for words embedded in sentences, either during reading or listening — diverges from most of the evidence reported in the literature.

I performed two brief searches in the literature available at the main database at Periódicos da CAPES. In the first one, I used the following keywords: cognate language activation. I have found a total of 214 studies. I examined the abstracts of all 214 studies and selected the ones which reported quasi-experimental linguistic investigations with adult participants. Among them, 72 studies had investigated cross-language activation via cognate words with adults: Allen, 2019; Allen, Conklin, Miwa, 2021; Anton, Dunabeitia, 2020; Boukrina, Hanson, Hanson, 2014; Branzi *et al*., 2020; Branzi, Martin, Biau, 2023; Bultena, Dijkstra, van Hell, 2014; Bultena, Dijkstra, van Hell, 2015; Cai *et al.*, 2011; Cassol Rigatti, Arêas da Luz Fontes, 2022; Chen *et al.*, 2022; Christoffels, Firk, Schiller, 2007; Comesaña et al., 2012; Comesaña *et al.*, 2021; De Bleser *et al*., 2003; Declerck, Özbakar, Kirk, Strijkers, 2021; Dijkstra, Grainger, van Heuven, 1999; Egan *et al.*, 2019; Fei, Zhao, Liu, 2022; Forcelini, Sunderman, 2020; Freeman, Blumenfeld, Marian, 2016; Freeman, Blumenfeld, Marian, 2017; Fricke, 2022; Gavino, Goldrick, 2020; Gullifer, Kroll, Dussias, 2013; Hameau, Biedermann, Nickels, 2021; Helms-Park, Perhan, 2016; Hoshino, Kroll, 2008; Hsieh *et al*., 2017; Hsieh *et al.*, 2021; Iniesta *et al.*, 2021; Jackson, Massaro, Hopp, 2017; Jacobs, Fricke, Kroll, 2016; Joss, Virtue, 2010; Kazemi *et al.*, 2023; Kennison, Fernandez, Bowers, 2013; Kim, Davis, 2003; Kootstra, Dijkstra, van Hell, 2020; Krogh, 2022; Lalor, Kirsner, 2001; Lauro, Schwartz, 2017; Lemhöfer, Dijkstra, 2004; Lemhöfer, Huestegge, Mulder, 2018; Lemhöfer, Spalek, Schriefers, 2008; Li, Gollan, 2018b; Li, Gollan, 2021; Lijewska, 2023; Lijewska, Chmiel, 2015; Lim, Christianson, 2023; Marcotte, Ansaldo, 2014; Martínez-García, 2019; Martinez-Garcia, Tremblay, 2016; Mishra, Singh, 2014; Mulder *et al*., 2014; Mulder, Dijkstra, Baayen, 2015; Poarch, van Hell, 2014; Poort, Rodd, 2019; Preuss, 2012; Schwartz, Kroll, 2006; Soares *et al*., 2018; Soares *et al*., 2019; Strijkers, Costa, Thierry, 2010; Szubko-Sitarek, 2011; Toassi, Mota, Teixeira, 2020; Tytus, 2019; van Assche *et al*., 2009; van Assche *et al.*, 2011; Wang, 2016; Von Grebmer Zu Wolfsthurn, Pablos, Schiller, 2021; Woumans, Clauws, Duyck, 2021; Xiong, Verdonschot, Tamaoka, 2020; Zhang *et al*., 2019.

Among these 72 studies, 52 presented evidences of cognate facilitation (Allen, 2019; Allen; Conklin; Miwa, 2021; Boukrina; Hanson; Hanson, 2014; Branzi *et al.*, 2020; Branzi; Martin; Biau, 2023; Bultena; Dijkstra; van Hell, 2014; Cai *et al.*, 2011; Cassol Rigatti; Arêas da Luz Fontes, 2022; Chen *et al.*, 2022; Christoffels; Firk; Schiller, 2007; Comesaña *et al.,* 2012; De Bleser *et al.*, 2003; Declerck *et al*., 2021; Dijkstra; Grainger; van Heuven, 1999; Fricke, 2022; Gavino; Goldrick, 2020; Gullifer; Kroll; Dussias, 2013; Iniesta *et al*., 2021; Jackson; Massaro; Hopp, 2017; Jacobs; Fricke; Kroll, 2016; Joss; Virtue, 2010; Kazemi *et al*., 2023; Kennison; Fernandez; Bowers, 2013; Kim; Davis, 2003; Kootstra; Dijkstra; van Hell, 2020; Lalor; Kirsner, 2001; Lauro; Schwartz, 2017; Lemhöfer; Dijkstra, 2004; Lemhöfer; Spalek; Schriefers, 2008; Li; Gollan, 2021; Lijewska; Chmiel, 2015; Lim; Christianson, 2023; Marcotte; Ansaldo, 2014; Martínez-García, 2019; Mishra; Singh, 2014; Mulder *et al*., 2014; Mulder; Dijkstra; Baayen, 2015; Poarch; van Hell, 2014; Poort; Rodd, 2019; Preuss, 2012; Schwartz; Kroll, 2007; Soares *et al*., 2018; Soares *et al.*, 2019; Strijkers; Costa; Thierry, 2010; Szubko-Sitarek, 2011; Toassi; Mota; Teixeira, 2020; van Assche et al., 2009; van Assche *et al.*, 2011; Wang, 2016; Woumans; Clauws; Duyck, 2021; Xiong; Verdonschot; Tamaoka, 2020; Zhang *et al*., 2019.), ten showed cognate inhibition (Anton; Dunabeitia, 2020; Bultena; Dijkstra; van Hell, 2015; Comesaña *et al*., 2021; Fei; Zhao; Liu, 2022; Forcelini; Sunderman, 2020; Hsieh *et al.*, 2021; Krogh, 2022; Lemhöfer; Huestegge; Mulder, 2018; Li; Gollan, 2018b; Martinez-Garcia; Tremblay, 2016), and five indicated no effects from cognate words (Comesaña *et al*., 2021; Egan *et al.*, 2019; Helms-Park; Perhan, 2016; Lijewska, 2023; Von Grebmer Zu Wolfsthurn; Pablos; Schiller, 2021). Also, one of these studies reported both inhibition and null effects from cognate words (Comesaña *et al.*, 2021), and two showed both inhibition and facilitation effects (Hameau; Biedermann; Nickels, 2021; Lemhöfer; Huestegge; Mulder, 2018). Two studies could not be accessed (Hsieh *et al*., 2017; Tytus, 2019).

In the second search, I used the following keywords: cognate language activation minority. I have found 7 studies investigating cross-language activation via cognate words whose participants were speakers of minority languages (Bosma *et al*., 2019; Bosma; Nota, 2020; Campos, 2023; Davis; Bowman; Kaushanskaya, 2018; Kirk *et al.*, 2018; Muntendam *et al*., 2022; Woolpert, 2019). Only two of these studies had adult participants speakers of minority languages or dialects and both showed cognate facilitation effects (Kirk *et al.*, 2018; Muntendam *et al*., 2022).

A Brazilian series of studies with minority language speakers did not appear in this search, but should be mentioned. Limberger (2018) observed that adult speakers of BP, standard German and Hunsrückisch — a German-based minority language — displayed a facilitation effect from cognate words in a generalized lexical decision task. A cognate facilitation effect was also seen for these speakers for a specific verb tense in a sentence comprehension task. Thus, the knowledge of Hunsrückisch influences the processing of German and the reading ability in German influences the processing of Hunsrückisch for these participants. Limberger (2018) verified that cognate words had an influence the processing of German and of Hunsrückisch, but not of BP.

Hypotheses 1, 2 and 3 in this study expected to observe a cognate facilitation effect in both lexical decisions tasks just like in Brenders, van Hell, and Dijkstra (2011), Jared *et al.* (2012), Schröter and Schroeder (2016), Bosma *et al.* (2019), Arêas da Luz Fontes *et al.* (2021), Bosma and Nota (2020), and Gastmann and Poarch (2022). Hypotheses 1 and 2 implied that lexical access for bilingual or multilingual people is non-selective. This was based on a vast literature on the bilingual lexicon and on predictions from bilingual language processing models. For example, BIA+ states that "[t]he information flow in bilingual lexical processing proceeds exclusively from the word identification system toward a task/decision system on the activation state of words" (Dijkstra, 2005, p. 197). This means that any top-down influences have no impact on the activation of words and their levels of representation. In addition, "cognate effects persist even when strong language membership cues are present in the absence of overlap between orthographies" (Winther; Matusevych; Pickering, 2023, p. 113). So even in a monolingual context, the bilingual person would still be affected by language co-activation.

Also, according to the BIA+ model, cognate words would be activated in parallel independently of whether the bilingual person was aware of form and meaning similarities between these words across languages. Davis, Bowman and Kaushanskaya (2018, p. S23) mentioned that "explicit training on how to utilise cognates in the service of reading comprehension" could be useful. Indeed, Hipfner-Boucher *et al.* (2016) showed that, since grade 1, children may demonstrate being aware of cognates. However, intervention for raising awareness of cognate words does not seem to help learning these words (Otwinowska *et al.*, 2020). Thus, the lexicons from both languages seem to be accessed at the same time via cognate words in a non-selective manner, and the awareness of cognate words does not appear to influence their processing.

Contrarily, results from this study are not in line with the majority of the existent body of evidence indicating that cognate words are accessed in a non-selective way. My results do not point to any facilitation effect and do not suggest any reliable interference effect from cognate words. Thus, these null results could be in accordance with a selective view of lexical access. Dijkstra (2005) explains null results in relation to two possibilities of phenomena. The first possibility would be that “the item types were not really comparable or were not matched properly” (Dijkstra, 2005, p. 182). The second one would be either that there was not enough relative activation of one of the languages or that participants would not always respond to according to task instructions (Dijkstra, 2005, p. 182). I will discuss the results from both studies considering these possibilities.

4.3.1 Discussing stimuli

There were two important limitations in terms of stimuli selection for this study. The first one was frequency of occurrence and the second one was length. There were also limitations in relation to the SC tasks, which will be addressed later in this subsection.

I could not find an annotated and normalized corpus of HC with annotations for parts of speech from which to extract the stimuli word. Instead, I compiled, translated, annotated and organized by estimated frequency HC words based on a HC-English dictionary and a HC-BP glossary. All these steps were guided and revised by an informant who was a highly-educated native speaker of HC who was living in Brazil since 2014 and who was enrolled in a Brazilian university course. The informant organized all 5,617 words according to his perceived frequency of occurrence. A total of 5,291 of them were categorized as being highly frequent, and this was the original set of words from which I selected cognate and noncognate words.

I understand that perception is a subjective measure and that perception of frequency does not equal actual frequency. I also understand that the frequency estimates originated from the perception of only one person. These frequency estimates would have been more reliable if they had been gathered from a bigger number of informants. I did find another comparable informant; however, they had very limited time available and could not contribute. Thus, it is clear that the frequency estimates in this study may be questioned.

The second limitation was word length. It was not possible to keep it constant due to the number of cognate words available from the final word set. Consequently, stimuli words varied in length from 3 to 12 characters, and participants’ RTs varied according to word length. Exploratory analyses (Tables 52 to 55) showed that longer words presented longer RTs in the LD task ($p < 0.001$) and in the WSC task ($p < 0.001$). There was also a marginal interaction between word length and cognate status in the LD task ($p = 0.0561$, unadjusted) which was not considered further due to its significance level. There was no effect of word length in the SSC task. Thus, as mentioned in Section 4.2.2, there was no significant main effect of length and no significant and reliable interaction between cognate status and word length. That is, the fact that longer words were also cognate did not facilitate (nor inhibit) their processing in comparison with noncognate ones. For this reason, I believe the fact that stimuli varied in word length did not obscure the cognate effect for this sample of participants.

Now I would like to address some errors and limitations of both of the sentence comprehension tasks. The first one is a programming mistake which was only detected after data had been analyzed. The second one is an error in the data output of the spoken SC task. And the third one is a limitation of the comprehension questions which was identified during data collection. The second and the third issues were compensated for with the same strategy.

The first mistake was already mentioned in Section 4.1.4.3. Sentences were presented in blocks of 5 so that the comprehension question always followed its respective sentence. So, the order of sentences within blocks was planned to be fixed while block presentation was randomized. However, when implementing this fixed order inside blocks, the implementation for mixing the cognate status of sentences inside blocks did not work. This means that each block only presented one type of sentence, either cognate or noncognate. This was not planned, but is an important flaw of the design of the SC tasks because the sequential presentation of 5 sentences containing cognate words could be noticed by a more skilled participant. Thus, any purely bottom-up effect expected to arise from cognate words (Dijkstra, 2005) could be biased by the awareness of the inclusion of cognate words in the tasks. Considering the literature, I would deem unreliable any language co-activation effect seen in the SC tasks after identifying this error. However, no significant effect was observed after correcting the p-values in neither SC task. In addition, the fact that SC tasks were self-paced might have compensated for this programming error. Since a sentence was presented on screen one word at a time, subsequent words might act as masks to previous words. It might be that the self-paced presentation style helped masking previous words in the task. Nonetheless, there is no possibility for solving this mistake for this study. Future studies must actively avoid this issue.

The second mistake is that, in the data output for the spoken SC task, accuracy for comprehension questions was not calculated because participants' answers for the question were not recorded. Accuracy was only calculated in the data output for the written SC task. The objective of the comprehension questions was to make sure participants were concentrated in understanding the sentences and were not distracted. This was compensated together with the third issue.

The third limitation was noticed during data collection with the group of adults invited at the church. A few of the participants asked about the meaning of words in the comprehension questions. They were explained that they should try and understand the meaning from the context. I took notes of these moments. It is possible that participants gave wrong answers to comprehension questions because they did not know the meaning of some words in BP. This means that the comprehension questions did not achieve their goal in the tasks and did not indicate whether participants were concentrated or not on the task. In order to compensate for these issues with the comprehension questions, I excluded from analyses the data from any participant who showed moments of distraction — such as looking repeatedly away from screen, talking with other participants — during data collection ($N = 7$) and used a more lenient threshold of 50% of correct answers in the WSC comprehension questions for including data from both SC tasks in the inferential analyses (Section 4.2.2).

4.3.2 Discussing language activation

In this section, I will consider the explanation for null results posited by (Dijkstra, 2005, p. 182) about the relative activation of one of the languages of participants. There was no indication of language co-activation either in the LD or the SSC or WSC tasks. It would be possible to argue that the lack of a cognate effect in the written tasks could stem from weak word decoding skills and vocabulary. However, if poor reading skill were the only reason for null results, cognate effects could have emerged in the spoken task. In this case, lack of vocabulary or linguistic knowledge could explain the absence of cognate effects. This might indicate that the parallel activation of languages was not visible in either RTs or accuracy rates due to more general language processing aspects.

Accuracy rates in the tasks varied vastly (Table 48). This meant that more than 50% of the data for both SC tasks and more than 25% of the data for the LD task was not added to inferential analyses. Considering all criteria used for including data in the final analyses, the variation in accuracy might be probably originated in participants' linguistic knowledge and language skills.

Now I will consider the information about the population sample in the present study which was collected via questionnaire. Participants who answered the language history questionnaire in full ($N = 25$) reported having almost excellent proficiency in HC (6.49) and good proficiency in BP (5.18). Their answers for hours using each language per day indicated much variation. For example, one person reported reading in HC for fun for 16 hours per day while another reported only 1 hour. One person reported watching tv in BP for 12 hours per day while another reported only 1 hour. Thus, most SDs for hours using the language are very high (Table 37).

Most participants in this study (*N* = 23) reported having a job. This characteristic is important for this sample because it was the determining factor for data collection. Participants had either long or strict working hours. For example, they would get up as early as 5am in the morning to commute to work and would manage to get back home after 8pm. This means that they probably had few opportunities to have conversations at home. Some participants added that their jobs would not require much talking during the day, which probably indicates few interactions in Brazilian Portuguese. Also, most participant mentioned that they read the Bible every day or almost every day in any of the languages they know.

Thus, although participants reported having good levels of proficiency in BP, this self-reported proficiency was not reflected in their accuracy rates. The self-perception on language abilities may be influenced by their communicative successes in everyday language use situations. However, these levels of perceived proficiency may not have been not sufficient to reach 70% of correct answers in the tasks completed in the studies or these levels may not equal knowing the vocabulary present in the tasks. In addition, the impact of proficiency on language co-activation and on the cognate effect seems to follow an inverted U-shaped model: in order for these effects to emerge, the bilingual person must be between a minimum level of proficiency in the L2 and a maximum one (Bultena; Dijkstra; van Hell, 2014). In other words, having too little knowledge of the language or having too much knowledge/proficiency will not produce co-activation effects. Therefore, one possibility for explaining the null results in this study is that there was not enough cross-language activation due to lack of linguistic knowledge in order to cognate words to have any impact on RTs or accuracy rates.

I was able to find one study investigating cross-language activation with an adult sample of speakers of a heritage/minority language. Muntendam *et al*. (2022) invited second-generation speakers of Turkish as an L1 and Dutch as an L2 — the dominant language — to perform auditory LD tasks in both languages. There were cognate and noncognate words across languages, and words also varied according to position of the stress syllable. There was no specific SES information for this group of bilinguals, but it is highlighted that all of them had completed at least secondary school. Results showed that, when executing the Turkish LD task, there was a general non-significant inhibition effect of cognate words in relation to noncognates, while when completing the Dutch LD task, the cognate effect was modulated by the position of the stress syllable. More specifically, there was inhibition when the stress syllable was the penultimate one and facilitation when it was the ultimate one.

These results are explained in terms of visual word processing and of a sequential language verification process in that participants would check whether a word belonged first to their L1, which is the strong one, and then to their L2, which may lead to inhibition or null effects. Muntendam *et al*. (2022) comment that participants had longer RTs in the Turkish task (around 1,000 ms) in relation to the Dutch one (around 890 ms), which may be caused by participant insecurities and double-checking for language membership of a word.

The results from this dissertation study are in line with the results from experiment 1 in Muntendam *et al*. (2022) in that there were no significant effects of cognate words. However, in Muntendam *et al*. (2022), the null effects appeared when participants were completing an auditory LD task in their L1. In the analyses I reported in Section 4.2.2, null effects appeared when participants were executing tasks in their L2. A relevant difference across population samples might be the language dominance. The participants in Muntendam *et al*. (2022) were dominant in their L2 while participants in my study seem to be dominant in their L1. For example, when answering what language was the most comfortable for speaking at home, 22 participants chose HC and 11 chose BP; when answering what language was the most comfortable for speaking with friends, 17 participants chose HC and 18 chose BP. Thus, participants with an L1 stronger than the L2 would be expected to show co-activation effects of the L1 when completing a task in the L2. However, this was not observed in any of the tasks in this study.

In the next section, I will summarize the results from Study 1 and 2 and offer a tentative explanation for the lack of cognate effects observed in this dissertation.

5 GENERAL DISCUSSION

The two studies reported here were guided by the following research questions:

1. Considering that bilingual and multilingual speakers have their languages co-activated at some level (Marian; Spivey, 2003; Thierry; Wu, 2007; Lemhöfer *et al.*, 2008), do HC-speakers show any influences from HC on lexical access during spoken and written comprehension in BP? More specifically, may HC co-activation facilitate BP processing?
2. Do BP phonological awareness and reading habits in BP play a role in comprehension of BP for HC-BP bilinguals?

Research question number 1 originated the general objective, while research question number 2 originated two specific objectives. The general objective was to investigate the effects of cross-linguistic interaction between HC and BP on lexical access during comprehension of spoken and written BP by native speakers of HC. The specific objectives were two-fold:

1. To investigate whether phonological awareness (PA) in BP influences the processing of BP as an L2 by native speakers of HC and whether it affects language co-activation effects;
2. To investigate whether reading habits (RH) in BP influences the processing of BP as an L2 by native speakers of HC and whether it affects language co-activation effects.

In order to pursue this general objective, I conducted two separate studies: one with children and teenagers — Study 1 —, and the other with adults — Study 2. Study 1 also examined the specific objective number 1, and Study 2 also examined the specific objective number 2. The reason for having two separate studies was to observe the cognate effect across different developmental stages, reading abilities, and language use characteristics. The goal was not to compare Study 1 and Study 2 due to participants age and language experience differences.

Results from Study 1 showed that children and teenager speakers of HC who migrated to Brazil do not appear to be influenced by cognate words across HC and BP when they are reading or listening to words in BP. However, the accuracy rates with which they decide whether a written word belongs to BP or not seems to be positively impacted by their level of PA. Results from Study 2 showed that adult speakers of HC who migrated to Brazil do not appear to be influenced by cognate words across HC and BP when they are reading or listening to words in isolation or in sentences in BP. Also, the frequency with which they read materials in BP does not seem to have any impact on their reading or listening to of words in BP.

Thus, results from the two studies reported here diverge from most of the evidence available in the literature on cross-language activation. The only expected result was the positive effect of PA in children's accuracy scores in a written LD task. All the other results were unexpected. In order to attempt to explain the unexpected null cognate effects presented here, I consider fifth possibilities. Four of them are limited alternatives and might be playing a role behind the results, but are not enough to clarify the big picture. The fifth one is my speculative interpretation of the observations from both studies.

The first possibility is related to L2 proficiency and attentional control during the task. Data collection occurred in a completely BP monolingual setting since I did not use any instance of HC with participants. Considering that the only language expected to be used was BP, it may be that participants were aware of the fact that they were not highly proficient in BP. Muntendam *et al.* (2022) mentioned that participants might have been insecure when deciding to which language the cognate words present in the tasks belonged, and this would not be implausible for participants in this dissertation. In addition, Bril, Gerrits, and Visser (2022, p. 270) remind us that "[l]ess proficient listeners [...] often show more controlled and conscious processing (Segalowitz, 2003)". Since participants in both studies were not highly proficient in BP, it might be that participants were so controlling and aware of their listening and reading skills during the tasks that any earlier, more automatic processes could not be visible in the response times. However, it is not possible to examine this potential influence at the time.

The second possibility is related to the amount of overlap of orthographic and phonological characteristics of the words used as stimuli across HC and BP. It has been shown that the cross-language activation of cognate words is modulated by the orthographic and phonological similarity their present, that is, the orthographic and phonological overlap. The higher the overlap, more facilitation is seen, especially in the case of identical cognates (Comesaña *et al*., 2015; Dijkstra; Grainger; van Heuven, 1999; Lemhöfer; Dijkstra, 2004). Any influences from cognate words could also have not appeared due to the variation in orthographic and phonological overlap across languages, which was not controlled here when selecting stimuli. This explanation may be tested in the future.

The third possibility is associated with conflict which is shown to be caused by cognate words during language production. Davis, Bowman, and Kaushanskaya (2018) mention this tendency both in behavioral and in neurophysiological results: "while behaviourally cognates were named faster than non-cognates, EEG data showed that bilingual speakers recruited domain general control operations when producing cognates" (p. 4). In both studies reported here, many participants were rehearsing the words overtly during the tasks, even in the auditory ones. They were never instructed to do so. On the contrary, I explained that they should silently answer as quickly as possible, so they were discouraged to read out-loud or repeat words. Overtly muttering the words could have slowed them down in that they would recruit more control in order to mutter the words. And, since "the cognate effect should appear early in processing" (Winther; Matusevych; Pickering, 2023, p. 112), it could be that this overt repetition masked any co-activation effects from cognate words at least for some participants. However, it is not possible to examine this potential influence at the time.

This conflict might be originated from word similarity across other languages known by the participants. In Study 1, other than HC, 27 participants reported also speaking French, 8 reported also speaking Spanish, and 2 reported also speaking both French and Spanish. In Study 2, other than HC, 24 participants reported also speaking French, 9 reported also speaking Spanish, and 9 reported also speaking both French and Spanish. In these cases, it is possible that cognate words may have activated words from French or Spanish due to some characteristics being shared among Latin languages. I could not produce a reliable strategy for separating words which were cognate across HC and BP but not across HC and French or Spanish. Thus, some participants could have displayed effects from the similar form of words from languages other than HC and BP. This possibility may be investigated in future studies.

The fourth alternative derives from the previous two. Cognate words have also been shown to cause interference, that is, slower RTs or lower accuracy rates (Dijkstra *et al.*, 2010; Comesaña *et al.*, 2015; Martinez-Garcia; Tremblay, 2016; Lemhöfer; Huestegge; Mulder, 2018; Forcelini; Sunderman, 2020; Hsieh *et al.*, 2021). Thus, it could be that the presence of cognate words slowed down lexical access due to their similarity between languages (Davis; Bowman; Kaushanskaya, 2018). However, this effect was not seen in any of the models run for any of the tasks. One possibility for explaining this is that cognate words could have confused participants so much (Muntendam *et al.*, 2022) that they took longer to make decisions about any of the words, including noncognates. That is, participants might be aware of a lack in lexical knowledge and might be confused by the presence of BP words which looked and sounded like HC ones. The parallel activation of lexical representations from both languages could have spread this activation through the semantic network, making other words available, increasing competition. Davis, Bowman, and Kaushanskaya (2018, p. 11) comment that "several studies suggest that cognate words induce response conflict (Acheson *et al.*, 2012), and that this response conflict may slow down language processing immediately after exposure to a cognate (Broersma *et al.*, 2016)". Nevertheless, these three speculations do not account for the absence of any type of cognate effect and for the variation in RTs and accuracy scores.

The fifth and last possibility considers a lack in lexical knowledge, which originates from poor experience with words. For this, we need to contemplate the following aspects. In Study 1, comments from some participants revealed that they did not engage much in conversation in any of the languages they knew. Also, most participants in Study 1 reported that they spend many hours per day using social media. Although social platforms allow for social networking and conversation, which may contribute for language learning, there is very little empirical studies about their efficacy (Lomicka; Lord, 2016, p. 258; Florit *et al.,* 2021, p. 15). In Study 2, Hypotheses 3 and 4 were not confirmed, which indicates that participants' RH did not influence their performance in any of the tasks. Also, few participants in Study 2 reported reading various types of materials, but many of them mentioned that they mostly read the Bible, which might indicate that they are exposed to a restricted vocabulary. In addition, general task accuracy in both studies was not in consonance with the self-reported proficiency levels, which is illustrated by the percentage of data removed from analyses for participants who had accuracy rates lower than 70%. In addition, the fact that cognate words did not have any influence on participants' RTs or accuracy scores might indicate that the bottom-up orthographic and phonological overlap was not enough to lead to co-activation of meaning across languages as is predicted by the BIA+. Thus, I would like to consider the impact of individual differences in relation to linguistic knowledge on cross-language activation in that this knowledge may also be defined by social aspects of the events experienced by the participants in the studies here reported.

5.1 LEXICAL KNOWLEDGE

Winther, Matusevych, and Pickering (2023) presented and compared two accounts for explaining the co-activation effects seen with cognate words. The first one is called online parallel activation and is usually the one used for interpreting co-activation effects. In the case of cognate words, the online parallel activation states that the semantic representation of both words is activated in parallel due to their shared orthographic and/or phonological characteristics. To put it simply, the semantic representation receives activation from two sources — two lexicons — and is activated at the same time for both languages, which facilitates lexical retrieval (Winther; Matusevych; Pickering, 2023, p. 107). This account is one of the assumptions of the BIA+ model, including more detailed influences from resting-level activation of words, which, as mentioned in Section 2.3.1, may be modified by proficiency and state of language activation in general (Dijkstra, 2005).

The second account is the learning account, which was originally proposed by Costa *et al.* (2017). According to this view, cognate effects would arise because of how these words were learned. More specifically, "the new acquired words will be structured according to semantic relationship, resembling the organization of the first language" (Costa *et al.*, 2017, p. 1633). Learning a word in a second language and associating their meanings creates a link between them. When the orthographic representation of a word is activated during language processing, this activation is carried over to the semantic representation of the word, which in turn spreads activation to the orthographic representation of the cognate word in the other language. This link between representations creates a trace from the word in one language to the word in the other, which is even stronger for cognate words.

This connection between lexical representations draws from Hebbian learning similarly to some connectionist models. When two neurons — or (patterns of) lexical representations — are activated at similar times, they are associated and create a link which will be strengthened every time they receive activation at the same time. The learning account proposes that this interconnection is the reason behind cognate effects for bilingual people. The phonological — and orthographic — similarities across words from different languages attract similar patterns. A comparable idea is present in the DevLexII model (Li; Zhao, 2013) when the semantic network is being formed.

Therefore, when learning an L2, the structure of the L1 lexicon would influence the organization of the L2 one. Consequently, this integrated lexicon would be different from the monolingual one (Costa *et al.,* 2017, p. 1635) and probably would be different from the multilingual one. The learning account is promising in that it allows us to consider that the way L2 vocabulary is learned after the consolidation of an L1 has an impact on the structure of connections in the bilingual lexicon.

When acknowledging both the online parallel activation and the learning account, they do not seem contradictory. Actually, they sound as if they could indeed complement one another.

> An on-line account explains cognate effects in terms of activation as a cognate is processed. In contrast, a learning account explains them in terms of the representation of cognate words in the mental lexicon. [...] Of course, a hybrid account is also possible (Winther; Matusevych; Pickering, 2023, p. 111).

Let us imagine the following scenario: a person is learning another language after the consolidation of their first one. This person would encounter cognate words either in a classroom or in an immersion setting. Since bottom-up information is sufficient for activating similar orthographic and/or phonological representations of L1 words, this person would produce this first association between the previously-known L1 word and the newly-learned L2 word. Then, every time either of those words receives activation, it also sends it to its counterpart in the other language, activating it in parallel. This process reinforces the connection between words both within and between languages. Later, activation spreads to other words that may be connected to only one of those two initial ones, even if they are not associated in meaning. For example, *train* and *ham* are not semantically related in English. However, Thierry and Wu (2007) showed that answering whether the pair was related or not was facilitated for Chinese-English bilinguals. The Chinese translations for *train* (*huo che*) and *ham* (*huo tui*) share the first character, *huo*. And pairs of words with this type of repetition in the irrelevant language for the task elicited smaller N400s than pairs with no repetition. This result was explained using the online parallel activation account, but Costa *et al.* (2017) state it is also plausible when the learning account is considered. This result by Thierry and Wu (2007) is exactly what the learning account model designed by Costa *et al.* (2017) was able to replicate.

Therefore, the structure of the L1 may be one of the factors which will influence the learning of an L2. Then, we should remember that "[p]rocessing may be affected by the age at which the two languages are acquired, level of proficiency, regularity of use, or similarity between the two languages" (Costa *et al.*, 2017, p. 1642). Due to this complex collection of influences, it is safe to add that this process will vary depending on what words the bilingual person is exposed to at what moments and contexts and how many times they are exposed to them until they are fully learned.

For example, Hipfner-Boucher *et al.* (2016) tested the level of awareness of cognate words across French and English for children in grades 1 and 2. They explain that this awareness was already documented in grades 4 and higher. Participants were English speakers immersed in French instruction contexts and performed the following tasks: cognate awareness, French reading comprehension, French word identification and receptive vocabulary in both languages. Before testing, participants were instructed on the concepts of cognate words, false friends and noncognates. Children showed awareness of cognate words in both grades 1 and 2, but answered above chance only in grade 2. Also, this awareness predicted L2 reading comprehension one year later. Hipfner-Boucher *et al.* (2016, p. 397) propose that these results illustrate the importance of lexical quality to reading comprehension because cognate awareness may reflect “a deep and rich lexicon”.

The LQH (Perfetti, 2007; Perfetti; Hart, 2001) posits that the knowledge about orthographic, semantic, morphosyntactic, and pragmatic information of a word is essential for language processing. Lexical quality is the specification or detailing of four word features, or constituents – orthography, phonology, grammar, and meaning –, and the strength of another constituent which binds the others together. In a high-quality lexical representation, all constituents must be precise, fully-specified and well-connected. Consequently, this representation will be coherent and stable in the lexicon and its retrieval will be effective and reliable. The LQH predicts that more skilled readers will have higher lexical quality representations in comparison with less skilled readers (Perfetti; Hart, 2001). This prediction may be extended to language processing in general.

Lexical quality can be increased as practice with words is intensified, that is, “[...] LQ depends on experience with words” (Perfetti, 2007, p. 365). Reading is a great activity for improving lexical quality because it gives the reader the opportunity of more encounters with a specific word and with less-frequent words. This contributes to the specification of constituents. In addition, high lexical quality representations are desirable for “they are responsible for automaticity (or at least efficiency) of word identification, which is what allows processing resources to be devoted to higher level comprehension” (Perfetti; Hart, 2001, p. 76). When the information from all word constituents converge, word identification is much more efficient and precise (Martin-Chang; Ouellette; Madden, 2014). Therefore, higher order cognitive functions may be allocated to performing more complex task demands.

These predictions were tested by Perfetti and Hart (2001) and Taylor and Perfetti (2016), among others. Perfetti and Hart (2001) had less and more skilled readers complete a task in which they would have to decide whether two words were related in meaning or not. However, homophone words were part of the stimuli list. For example, participants would respond *yes* to *king* and *royalty*, which are indeed related in meaning. But they would respond *no* to *evening* and *royalty*, which are not semantically associated. Additionally, they would be presented with *night*, a homophone of *knight*, and *royalty*, which are also not related, yet were expected to cause confusion because of the phonological overlap. Perfetti and Hart (2001) observed that more skilled readers both noticed and solved the ambiguity of the homophones faster than less skilled readers. However, the homophone interference effect was not seen when sentence context was available, which may be explained by the presence of contextual information.

Taylor and Perfetti (2016) conducted a paragraph reading experiment and a lexical knowledge training experiment. In the first one, they observed that when participants had less reading experience, they took longer to read low-frequency words in comparison with higher-frequency ones. In the second experiment, considering their model analysis, when there was only word form information (orthography and phonology) available, eye movement measures were slowed down, but when semantic information was added, more experienced readers showed fewer word re-reading. In addition, less experienced readers were less benefited from the same number of word encounters than more experienced ones.

LQH predictions also hold for bilingual people. Raudszus, Segers, and Verhoeven (2018) demonstrated that lexical quality is a direct predictor of reading comprehension for both monolingual (Dutch) and bilingual (many L1s-Dutch) children. Participants were tested in Dutch for vocabulary, decoding, syntactic integration, working memory, inhibition, and reading comprehension. In their models, vocabulary, decoding, and syntactic integration had a direct impact on Dutch reading comprehension, while working memory and inhibition influenced syntactic integration. Additionally, they found that L1 vocabulary had an impact on L2 (Dutch) reading, which suggests that a richer vocabulary in the L1 may contribute to learning new words and to acquiring an L2. The study by Lervåg and Aukrust (2010) derived similar results. Decoding and vocabulary were essential in predicting reading comprehension in grade 1, but only vocabulary was able to anticipate reading comprehension development 18 months later. This was especially strong for L2 language learners. In addition, Verhoeven, Voeten, and Vermeer (2019) conducted a longitudinal study to test for the impact of lexical quality on reading comprehension of monolingual and bilingual children. It indicated that reading comprehension was predicted directly by morphological knowledge and receptive vocabulary, which were measured two years earlier.

Therefore, going back to the results of the two studies reported here, it may be that participants lack so much lexical knowledge in either language that most words can be a source of confusion, and no language co-activation effect arises. Perfetti and Hart (2001, p. 80) did observe a similar situation with homophone words: "Less skilled readers are confused by *gate*; they are probably confused by *gait* as well; but they are also confused by *stride*, a low-frequency control word, and thus have little opportunity to show a confusion that is specific to homophones".

The confusion observed by Perfetti and Hart (2001) in less skilled readers could be compared to the accuracy scores from participants in both studies. Their experience with words is not as varied and rich as the one from a skilled reader. And since this practice with words does not affect only reading skills, but language skills in general, performance in both task modalities — visual and auditory — was disturbed. Thus, the origin of the absence of a language co-activation effect does not seem to be decoding skills, but language skills more generally.

Although participants reported good levels of proficiency in BP, their lexical knowledge appears to fall behind it, and this is well-illustrated by the fact that more than half of the collected data was associated with accuracy scores lower than 70%. Most of the participants reported few opportunities or little motivation for practicing language skills. And when these moments were indeed mentioned, the variety of types of materials for listening and reading was not vast. Thus, it is plausible to question whether participants were able to create actually high-quality representations of the words they encountered in their everyday activities. They may be proficient enough to communicate with relative success in everyday social situations even with unstable lexical representations in their L2.

The fact that part of the participants mentioned to have few opportunities or motivation to engage in conversation in BP may be considered simply anecdotal evidence. However, it is not implausible in light of studies on the social inclusion of immigrant groups. A recent study shows that "low-SES and female immigrants appear to be doubly vulnerable in class, which specifically affects their friendships" and that "proficiency in the national language predicts the social inclusion of immigrant children, while controlling for all the other variables considered" (Cavicchiolo *et al.*, 2023, p. 146). This highlights the importance of experienced use of the majority language for children and adults, which also depends on those opportunities for conversation and for encountering new words and sentence structures.

After word decoding skills are automatized, the reader can concentrate higher-order cognitive skills on comprehension. At this stage, listening comprehension becomes an even more important factor for reading (Verhoeven; van Leeuwe, 2012). In this manner, listening comprehension is an essential skill for reading, together with decoding, as explained by Gough and Tunmer (1986, p. 7). The quality of experience with words — both written and spoken experience — may be the key for the results reported here in relation to task performance and the lack of a cognate effect. This is also in accordance with assumptions from the BIA and BIA+ models. They state that subjective word frequency and recency of use define the resting-level activation of words in each lexicon (Dijkstra, 2005, p. 190). In other words, how recently and how many times the bilingual person has encountered and used a word will set its activation parameters.

While lexical quality is an internal characteristic of lexical representations, it is also regulated by external factors which influence the linguistic input the person will be exposed to. In their study about text-reading with monolingual and bilingual children, Davis, Bowman, and Kaushanskaya (2018) observed that cognate words caused interference for bilinguals. They posited that, "due to the differences between groups in socio-economic status, language experience is unlikely to be the only explanation for lower reading skills observed for the bilingual children" (p. S22). And poverty does have serious impact on cognitive development (Storrs, 2017), language included.

A poor linguistic environment is shown to affect and to hold back children's literacy development, but it also impacts language development in general (Ece Demir-Lira *et al.*, 2019; Rigatti *et al.*, 2023). More specifically,

> the process of language acquisition is influenced by important environmental factors. It is known that in a family of readers there is a qualitative and quantitative difference in interaction and linguistic stimuli given to the child; this difference is echoed as qualitatively and quantitatively higher-quality interaction of parents and caretakers with their children. One of the results is that children in these families listen to 30 million words more than children living in families of lower economic and educational levels [...]. The result of this is that the child acquires a richer and larger vocabulary (Buchweitz; Mota; Name, 2018, p. 123).

Poverty is usually measured using surveys (The World Bank, 2024), and SES is one possible measure. "[S]ome neuropsychological functions seem more sensitive to SES effects than other: this is the case of language, memory, and executive functions" (Puglisi; Salles, 2018, p. 224). Puglisi and Salles (2018) explain that the development of executive functions takes longer than other aspects, reaching relative maturity from 25 to 30 years of age (Lebel *et al.*, 2008). In addition, environmental impacts during this initial developmental phase may have consequences which would be arduous to overcome later. Thus, poverty in early childhood is associated with low cognitive performance (Puglisi; Salles, 2018, p. 224).

More specifically, SES is associated with language development. As examples, Puglisi and Salles (2018) present two studies one the power of SES of explaining variability in language performance. The first one was carried out by Noble, McCandliss, and Farah (2007). A total of 150 children in grade 1 and from diverse SES backgrounds completed tasks involving language, memory, working memory, cognitive control, visuospatial skills, and reward processing. SES accounted for 30% of variability in the language tasks, which tapped lexical-semantic knowledge and the ability to combine phonemes to form words. SES also explained variability of the other constructs, but in smaller magnitudes (< 17%). In addition, regression analyses showed that "school environment partially mediates the association between SES and language abilities" (Noble; McCandliss; Farah, 2007, p. 475).

The second study was a Brazilian multicentric one, which was carried out in 17 public and private schools in São Paulo and in Salvador (Engel de Abreu *et al*., 2015). Its objective was to examine the effects of SES on children's cognition, including language. Academic achievement was tested with tasks involving oral language, reading, writing, mathematics, and sciences, while language ability was tested with comprehension and production tasks. Results showed that SES accounted for 50% of variability in tasks tapping language. Engel de Abreu *et al.* (2015, p. 13) emphasize the fact that in this study SES explains more variability than in studies conducted in countries with higher Human Development Index (HDI) and posit the following interpretation. On the one hand, when comparing families of medium and high SES, they can offer a more consistent range of cognitive resources to the children, that is, there is less variation in the resources offered to children by medium- and high-SES families. On the other hand, when comparing families of medium and low SES, variability of the cognitive resources which can be offered to the children increases. Thus, SES becomes a stronger predictor of cognitive — and linguistic — development in lower-SES families.

Especially during early childhood, poverty and SES have an intense impact in development. "Language is one of the most susceptible functions to socioeconomic effects, and during school age vocabulary differences between medium- and low-SES children end up being larger than in other areas, such as memory and attention" (Puglisi; Salles, 2018, p. 227). SES and other social and environmental factors are usually demonstrated to have indirect effects on language and cognition. However, these effects are also shown to predict later skills and academic achievement. For example, Corso *et al.* (2016) observed that SES explained 31% of variance in children's executive functions, which in turn accounted for 52% of variance in reading comprehension for texts. The SES effect was mediated by executive functions and had an important impact on reading skill.

There is another Brazilian study worth mentioning here. Piccolo *et al.* (2013) conducted a longitudinal study with mothers and their children of low SES background for over the span of ten years. Psychosocial tests were administered at three timepoints, and reading performance was tested when children were 9-10 years old. Results showed that family income was positively correlated with performance in the reading comprehension task. Also, the number of people living in the house with the child was the only predictor of word reading, and this was a negative relationship, that is, the more people lived in the house, the worse was the performance in the reading task.

These studies, among others, show that the environment is able to define an important part of language development. Although the ability to develop language is innate, exposure to linguistic input — especially rich and varied input — is essential for it to flourish. The focus of the studies just mentioned is on early childhood, but it would not be disparate to extend some of this indirect effect of SES on language to late L2 learners and adulthood. At least, it does impact motivation (Kormos; Kiddle, 2013). And this may be the reason for only one of our ten hypotheses having been confirmed.

Participants reported good levels of proficiency, which did not match general accuracy scores. No language co-activation and no L2 RH effects were seen. However, in Study 1, L2 PA did associate positively with accuracy in the written LD task. It is probable that participants' proficiency levels are not as good as they reported them to be and, consequently, that they lack high-quality lexical representations. In addition, it is possible that they were never exposed to the cognate words in the tasks previously, thus never triggering the steps proposed by the learning account. Also, the resting-level activation for cognate words may be so low that activation never reached a threshold, which would otherwise lead to parallel activation of words in both languages. Therefore, neither a facilitation nor an interference effect was seen.

High-quality lexical representations have fully-specified and well-bound constituents. This implies a richer and more varied experience with words. The consequence of high lexical quality is faster, more efficient and more automatic word decoding. Moving these ideas closer to a bilingual language processing model such as the BIA+, faster and reliable word decoding will spread activation faster as well since there is information for all lexical constituents of the word. It is plausible to expect that when all constituents are fully-specified, they will receive and send activation to the network more efficiently than when only a few constituents are available or when they are unstable. Few encounters with a word in only one modality, setting or context will not create a lexical representation as rich and varied as the one for a word encountered several times in different modalities, settings and contexts. Few encounters also mean a less strong connection between words or levels of representation, which is important for both the learning account and the lexical quality hypothesis.

It has been shown that highly balanced and proficient bilinguals do not present language co-activation effects in comparison with less balanced or proficient ones (Arêas da Luz Fontes; Schwartz, 2015; Cop *et al.*, 2016). And, just like in Perfetti and Hart's second experiment (2001, p. 79), more skilled readers did not display any homophone effect. Higher quality of representations speeds up lower-level processing in order to surrender resources for higher-level ones. Then, the speaker is free to deal with the activation of lexical competitors more efficiently. Nevertheless, when lexical quality is low and connections between representations are weak, the speaker will need to use more cognitive resources to finish this step before solving any conflict in language co-activation.

In summary, when it comes to linguistic aspects, the experience with words should be as rich and varied as possible in order for language processing to be as efficient as possible. So, if it is indeed the case that the word experience for the participants in the studies reported here has not been as rich and as varied as the ideal one would be, why is it so? I believe it stems from social and environmental factors, such as SES.

For example, Arêas da Luz Fontes *et al.* (2021) did observe the cognate facilitation effect when BP-English children performed a lexical decision task in their L2. However, it is necessary to highlight that the participants in that study were students at a highly prestigious bilingual school in Brazil. This fact probably suggests that students came from higher-SES family backgrounds. This is an important demographic distinction between the study by Arêas da Luz Fontes *et al.* (2021) and the ones in this dissertation. There is no characteristic specific to the language pair HC-BP which would make me believe language co-activation would not be observed. This phenomenon has been observed even across languages with completely different writing systems (Thierry; Wu, 2007). In this dissertation, participants have migrated because of a tragic natural disaster and found themselves in lower-SES life conditions than the ones in Arêas da Luz Fontes *et al.* (2021). Participants in Woolpert's (2018) study were also in a context of minority and majority languages, belonged to migrant families and did not show any cognate effect. However, there is no information about SES for the participants in Muntendam *et al*. (2022), only that they were second-generation immigrants. In Study 1, no participant reported being born in Brazil. In Study 2, all participants had migrated to Brazil.

Bosma and Nota (2020) also tested the cognate effect in bilingual minority language speakers. Children were less dominant in Frisian, which was the minority language they spoke at home, and were dominant in Dutch. Cognate facilitation effects were only seen for identical cognates when children performed the sentence reading task in Frisian. There was no effect in Dutch, and non-identical cognates did not facilitate reading. Bosma and Nota (2020, p. 6) did report the SES of participants' backgrounds, which ranged from medium to high.

SES is not calculated only in relation to income, but also to the level of formal education of family members (APA, 2017). Although we do not have information about participants' income, we do have information about the family members' educational attainment (Tables 6 and 38). In Study 1, participants who completed the LHQ reported that 41% of mothers had finished elementary school, 43% finished high school, and 9% had studied at a university, and that 29% of fathers had finished elementary school, 56% finished high school, and 8% had studied at a university. In Study 2, participants who completed the LHQ reported that 21% of them had finished elementary school, 57% had finished high school, and 21% had studied at a university; that 20% of mothers had only finished primary school, 41% had finished elementary school, 29% had finished high school, and 8% had studied at a university, and that 8% of fathers had only finished primary school, 52% had finished elementary school, 21% had finished high school, and 17% had studied at a university. Higher education is usually associated with higher-SES backgrounds. Thus, it is possible to observe that our sample mostly consisted of low- and medium-SES families.

Another illustrative fact about our participants' social realities is where they were living. This is worth highlighting because it was one of the obstacles encountered during data collection. In 2015, the estimated number of Haitian immigrants living in Florianópolis was 8 thousand people (NSC, 2015). Part of the city of Florianópolis is on an island and part on continental land. During the beginning stages of data collection, I could not find HC speakers in most schools on the island. I was able to find them when contacting schools on the continent and in neighboring cities. Then, I received the information that the cost of living in Florianópolis, especially on the island, was too high for some families, and they ended up moving to more financially convenient places. This suggests that SES of those families may be low, at least at that point in time.

Finally, it is relevant to reiterate that the circumstances of migration for the population of interest here are not favorable. There are natural disasters which contributed to intensifying a previously delicate political and economic situation. Learning another language — the majority language — in this context is not a voluntary choice, it is a necessity in order to be included in society. Immigrants from Haiti, and from other countries as well, came to Brazil fleeing a social and political crisis and/or looking for work and better quality of life. For most of them, the routes taken to enter Brazil are not safe options, and they end up settling in the periphery of cities and are offered low-paying and physically demanding jobs (Butikofer; Silva, 2021).

These circumstances may be exacerbated by other obstacles. For example, difficulty in learning and using the language was the most commented issue Haitian immigrants found in 2013 (Sá; Silva, 2016, p. 5). The majority language is the key for better job opportunities and being part of the society. Not being able to communicate in the majority language hinders social mobility and exposes the migrant to situations associated with stereotypes and discrimination (Sá; Silva, 2016, p. 10). Lopez (2018) interviewed immigrant students and reported their perceptions and experiences. For example, they understand that not learning BP equals "the impossibility of performing several critical activities" (Lopez, 2015, p. 405) and that "knowing Portuguese implies a guaranteed access to certain activities and solutions for their problems" (Lopez, 2015, p. 406). Language knowledge is also seen as a self-defense tool because it is essential for exercising civil and human rights and for avoiding labor exploitation (Lopez, 2018). Thus, linguistic knowledge becomes not only a path for social inclusion but also one of the goals of the minority-language speaker.

Florit *et al.* (2021) conducted a six-month longitudinal study with monolingual and minority-language bilingual toddlers, all from low-income backgrounds. They investigated whether the home language activities, such as book reading, singing and storytelling, with interaction with the adults or with digital media would be associated with expressive vocabulary. Results showed that both groups offer similar duration and frequency of home language activities to toddlers, irrespective of language. So, belonging to a minority language context by itself does not imply that linguistic input will be less frequent than in a monolingual context. In addition, Florit *et al.* (2021) observed that, 6 months after the beginning of the study, total expressive vocabulary of toddlers was accounted for the frequency and duration of home language activities in which adults interacted with the children. The same relationship was not seen when activities involved passive exposure or digital media. This study highlights the impact of language use at home and of linguistic interaction for the development of any language.

Although being an immigrant in itself does not seem to influence language development (Florit *et al.,* 2021), it might be associated with factors which do so. Ibáñez-Alfonso *et al.* (2021) tested native and immigrant children on reading comprehension, sentence comprehension monitoring, syntax, and vocabulary in Spanish. They observed that the type of group did not explain variability at all, but SES, age, non-verbal reasoning and comprehension monitoring did. Despite there being no significant impact from type of group, "[t]he group of native children could be considered middle/low-SES, while both immigrant groups could be considered low-SES" (Ibáñez-Alfonso *et al.,* 2021, p. 4). It might be the case that some immigrants are at risk of ending up in a lower-SES context than non-immigrants.

The main idea here is that the score on a language proficiency test or the label of speaker of a specific language is not enough to guarantee that linguistic phenomena will be observed. Multilingualism is a complex continuum of experiences, and this may be the time to "rethink experience" in psycholinguistic studies as Titone and Tiv (2023) propose. They urge fellow psycholinguists to consider in their research the social factors which are known to influence language experience and development. Titone and Tiv's (2023) goal is to motivate us to consider how the experiences of individuals, of people interacting with other people and of communities and societies impacts language and cognition and to rethink definitions and approaches to languages and language use in relation to people's experiences and sociocultural characteristics. In order to do so, they present a Systems Framework of Bilingualism.

The Systems Framework of Bilingualism is inspired by social, interpersonal, ecological, linguistic and cognitive approaches in order to offer a more domain-general account. This framework combines four levels or layers: the interpersonal, that is, individual to individual, the ecological, that is, small communities and groups of people, the societal, that is, large societies and their values, and the temporal, that is, the natural development of the individual and historical time. These levels configure a complex system, in the middle of which the individual is located. One level contains and includes the previous ones: the individual is included in the interpersonal level, which is included in the ecological level, which is included in the societal level, which is included in the temporal level. All levels of the system influence the individual in varied ways and shape their cognition, language and neuroplasticity.

The Systems Framework of Bilingualism allows us to ponder about the complex dynamics of multilingualism which are investigated in detail when it comes to cognition and neurobiology but consider very little of the influences of social factors in the case of adult people. Titone and Tiv (2023) present evidence of these influences for the first three levels and they argue that "these ambient, contextual effects of bilinguals' lived social realities really have consequential and observable impact on BEHAVIORS themselves" (Titone; Tiv, 2023, p. 10, emphasis in original). Therefore, according to the account proposed in the Systems Framework of Bilingualism, it is plausible to conclude that the observed effect of cross-language activation in bilingual and multilingual people may also be modulated by the environmental, linguistic and social experience these bilinguals and multilinguals have in their every lives. Controlling and accounting for these variables in this complex system should be the most arduous task for psycholinguists who are interested in accepting the suggestion. However, the role of these external factors in bilingual and multilingual language processing may be the key to avoid missing "crucial discoveries about how bilingual or multilingual experiences enrich, or are enriched by a variety of in-the-wild social and cultural experiences" (Titone; Tiv, 2023, p. 12).

Finally, language learning is influenced by several factors which interact with one another. It is a challenging task to try and disentangle them in order to understand both details and the bigger picture. This challenge comes not only from the limitations in the research design, but also from the complex relationships among these constructs. Thus, this discussion is not made up of clear-cut statements, but of ideas which may agree with one another in the larger debate in the research field. I hope this tentative explanation may inspire future research questions. Lastly, I would like to summarize my interpretation of the results observed in this work: Language co-activation may not be seen in minority-language bilinguals from low-SES backgrounds because the environment will not provide language and word experience rich and varied enough for these bilinguals to perform efficient lexical processing in another language.

6 DATA COLLECTION EXPERIENCE REPORT

This is not a systematic report on the data collection method used in this dissertation. This is a report about the steps I took to conduct data collection and the difficulties I encountered. The process of data collection started in 2021 and ended in 2022. A survey was conducted from September 2021 to August 2022. Approval from the Ethics Committee was received on September 14th 2021 (CAAE 51539821.3.0000.0121). Meetings to invite participants started in February 2022 and the actual data collection started in August 2022 and ended in December 2022.

6.1 STUDY 1

Firstly, a survey was conducted to verify what schools had HC speaking students enrolled. I needed a list of schools where I would like to invite students so that I could ask the City and the State Department of Education for permission to enter schools. The information of the enrollment of HC-speaking students in the schools was needed to create a list of schools of interest. This step was mainly carried out by telephone calls. Emails were also sent, but very few were answered. School information, such as telephone numbers and e-mail addresses, were found in official City Hall and State Government.

I contacted schools and introduced myself very briefly. At first, I asked the following question: are there any students who speak HC in the school? However, soon I changed it to: are there any students from HC families in the school? This change was due to the fact that some schools did not have information about languages spoken by students, but they had information about countries of origin.

If the staff person did not have that information, I asked who I could talk to about this. Sometimes the person in charge was not available and it was necessary to call again. Other times they did not have that information at hand and requested to be called at another moment. Also, there were times when school staff reported to not have this information at all. In most cases, I needed to call schools more than two times to receive a rough estimate of how many HC-speaking students were enrolled. Few schools refused to answer this question via phone call. The ones which did not share this information via phone call requested that I visit the school in person and send documents regarding the research project. In many cases, I could only draw an estimate of the number of HC-speaking students when visiting the school.

In this research study, schools would have the role of mediators between the researcher and the students' families, so the staff did not share phone numbers from families. When I first visited the schools, I took paper invitations to be delivered to students. The invitation had text in BP and in HC (Appendix H). However, most schools did not have staff available to guarantee that invitations were handed in to students or to keep track of the families that would send back an answer to the school regarding their children's participation in the study. Some schools did not allow me to talk directly to students to invite them, while others allowed it. In some schools, staff organization seemed to be coordinated and active, and students' data seemed to be well cataloged; in these cases, checking the potential participants who were invited and what families responded (positively or negatively) to the invitation was much easier. However, most schools could not have one person in the staff to oversee this process since schools had their own community demands to manage. Thus, it took many in-person visits to schools until I had a list of potential participants.

Many times, I had to revisit the school because students who would be invited had not gone to class that day. There were times when I asked that students would be reminded of my visit so that they would go to school; however, they did not show up. When this happened — that is, when students would miss many school days or would not give the invitation or the TCLE back —, I did not insist on inviting them anymore. This decision was made because some Haitian families were a bit weary of new projects. In addition, three schools reported that students mentioned that they would soon move to another country, such as Canada or the USA, and that their families would part on foot. Staff did not know if these comments were true or if they would actually happen, but they mentioned them in concern for the students. Moreover, there were times when students changed schools or dropped out.

When potential participants were listed, I printed TCLEs to be sent to families. Some families answered right away, allowing or denying participation. In other cases, TCLEs had to be sent twice or more until an answer was given. At first, the researcher delivered only the TCLEs to schools to hand them in to students. However, since some families took many weeks to give them back, I started to distribute TCLEs together with the language history questionnaire (LHQ), explaining that if families allowed their children to participate, then they would fill the questionnaire. There were four cases in which families answered less than one third of LHQ questions and did so using the wrong scale (for example, the question would be "how old were your children when they started to speak HC?", but the answer would be "yes"). In these cases, I asked the school staff or the children whether they knew the answers; if they did not, the question was left blank.

At first, the population of interest was children from 6 to 9 years old enrolled in public schools in Florianópolis. However, the total number of potential participants found in the survey did not reach the expected minimum, which was 40 participants. Thus, I sent a request to the Ethics Committee to broaden the area of search for participants and the age range to include older children and teenagers. Then, the final group of participants was from 7 to 17 years old and was enrolled in public schools in Florianópolis, Palhoça and São José.

From February to September 2022, I could reach schools by car due to my family and my colleague's kindness. From September to December 2022, I mostly used public transportation. Mean transportation time from home to data collection site was 2,5 hours.

The day for participation was scheduled with the school staff since most families and teachers allowed the students to step out of class to participate. There were only two cases when teenagers agreed to come back to school in the afternoon, when they did not have classes, to participate. Participation occurred in different rooms across schools: teachers' room, supervisor room, library, presentation room, computer room, empty classrooms. Noise level varied across schools and periods of the day; it was very rare to find a quiet place for data collection.

Participation sessions followed the script displayed in the data collection guide (Appendix K). They always started with an explanation of what would happen in the session and of the contents of the Willingness Term, which is a term for underage participants to indicate that they are willing to participate in the study, that they are voluntary participants. If participants were younger than 11 years of age, participation started with the letter identification task; if they were older than 11, it started with the lexical decision tasks. The LHQ was answered by families if the student was younger than 11 and answered by the students themselves if they felt comfortable with that and if they were older than 11. There were two LHQ versions, both printed: one in BP and another in HC. Most families answered it in Portuguese, but some opted for the HC version.

Data collection used mainly the researcher's personal laptop (described in Section 3.1.6) since most schools could not offer both a computer and a quiet place for participation. Usually, they did not have a computer available at all due to technical issues or full schedule of the computer room. Only twice a school computer was used; in this case, it was not possible to control technical specifications. Machines were described in Section 3.1.6. A colleague's laptop was also used and is described in Section 3.1.6. All participants used headphones or earplugs while performing the tasks; only one chose to not use headphones. Usually, participants used the researcher's headphones, but three participants had their own earphones and preferred to use theirs. Headphones mostly used were described in Section 3.1.6. Figures 63, 64 and 65 illustrate how data collection was conducted.

Figure 63 – Participant in Study 1 completing the phonological awareness task

Source: The Author (2024)

Figure 64 – Participant in Study 1 completing the visual lexical decision task

Source: The Author (2024)

Figure 65 – Participant in Study 1 completing the phonological awareness task

Source: The Author (2024)

I contacted 36 Florianópolis City public schools and 76 Santa Catarina State public schools, a total of 112 schools. I visited 5 Florianópolis City schools and 31 Santa Catarina State schools, a total of 36 schools. Finally, data collection actually took place in 2 Florianópolis City schools and 18 Santa Catarina State schools, a total of 20 schools.

6.2 STUDY 2

Adults were invited mainly online. The invitation had text in BP and in HC (Appendix I). At first, I invited adult participants for Study 2 by sending online invitations to Portuguese as a Welcoming Language Project (Projeto Português como Língua de Acolhimento, PLAM) at UFSC and to Brazilian Portuguese for Humanitarian Migration Project (Português Brasileiro para Migração Humanitária, PBMIH) at UFPR. Since there are university students who speak HC and since there are extension language classes for immigrants at UFSC and UFPR, I asked the departments to forward invitations to those potential participants. The invitation was a poster in BP and HC (Appendix H). It explicitly invited HC speakers to participate and invited those who would be interested to contact the researcher. There was a cellphone number and an e-mail address. However, no one answered back. I also shared the invitation in social media, such as Instagram, Twitter, and Facebook, especially in groups created by and for Haitian immigrants in Florianópolis. However, no one answered either.

Then, I asked colleagues about friends or friends of friends or students who would know any potential participants. One colleague mentioned that, after inviting a group of Haitian acquaintances, they apprehensively refused to participate. I concluded that maybe an in-person invitation approach would work better than a distant online one. Thus, I contacted schools from Study 1 which also offered basic education classes to young adults and adults (EJA). In one of them, staff indicated 10 potential participants who studied at night. I visited the school at night and met one of those students. I handed him some invitations, and he offered to deliver them to the others. However, he was the only one interested in participating.

This participant mentioned that the church he attended in Palhoça also received many other Haitian immigrants and he offered to invite them as well. He shared their preacher's phone number, and I talked to the preacher through WhatsApp messages. The preacher was thrilled to help with the study, and a visit to the church was scheduled.

The visit had two objectives: (1) to invite people in-person so that they would feel less apprehensive and be able to answer any questions, and (2) to examine the physical space of the church in order to check whether there would be room, furniture and internet connection to have data collection to happen there. The possibility of collecting data at the church was suggested because potential participants would work every weekday, from early morning to late evening, and would not be able to go to UFSC at night during the week to participate. In this case, laptops would be rented and taken to the church so that everyone interested would be able to participate. However, the visit showed that the physical space was not appropriate for data collection: there were no tables, only chairs, and internet connection was very unstable. In addition, prices to rent laptops for one day were too expensive for the research budget.

Once a week on weekdays, service was held at night. I went there by taxi accompanied by a fellow researcher who voluntarily assisted during data collection; the drive from Florianópolis downtown to the church, in Palhoça, lasted one entire hour. The preacher welcomed us and asked us to wait until service was finished. He spoke French when preaching, but also used HC at times. The preacher ended that specific religious moment and, after everyone sang collectively, he asked us to come near him and explain the reason for the visit. Around 40 adult people were present. In sum, I introduced myself and my colleague, explained the objective of the study and mentioned that it is one of the first to look into this pair of languages; my colleague added that participation would include reading in BP on a laptop or computer and told them it had been difficult to find people interested in participating. The preacher translated everything to French and asked who would be interested; most people raised their hands. Then, two people had questions about certification for participating and about the place where participation would occur. It was not possible to offer them any certificate for participating in the study and the place was yet to be confirmed, but it would probably be at UFSC. We thanked everyone profusely and got a lift back to Florianópolis.

It was agreed that the preacher would write down the names of the people interested in participating and send them to me; I had a blank list for this situation. Moreover, I would verify the possibility of taking participants to the university on a Saturday so they could participate using the computers in the technology room at the Center for Communication and Expression at UFSC.

The technology room was scheduled for September 10th from 7 pm to 10 pm according to the times the preacher mentioned would be comfortable for participants. The researcher was there 2 hours before to turn computers on, check earphone plugs and internet connection and to access the website where the experiment was hosted. The room was scheduled once more for September 24th, but earlier from 5 pm to 10 pm.

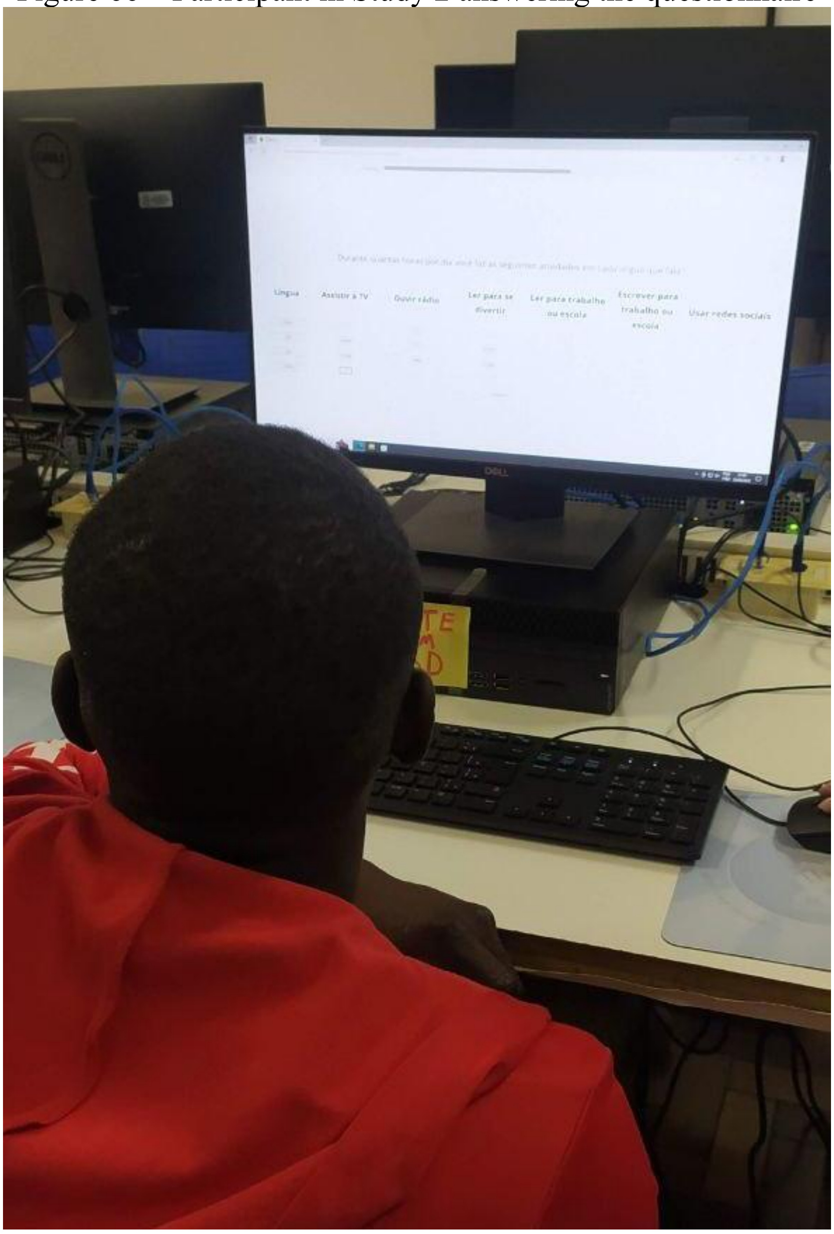
Participants were taken from the church to UFSC by van. The van was scheduled two days prior to the day of participation. The first van took 18 people to the university on September 10th; the second one took 10 people on September 24th. Two participants on September 24th were the preacher's children, so they completed tasks from Study 1. Van reservation was paid using funding from my supervisor's taxa de bancada (process number 311632/2019-0).

Twenty people were expected for each data collection day. Many moments of miscommunication and misinterpretation happened. For example, the preacher had understood that the first day of participation would occur on that very week when I visited the church. However, it was not possible at all, because the visit was imperative to decide whether a van would be scheduled or not. There were also misunderstandings concerning the time of pick up at the church. Two people were late on the first day and many more did not show up on the second one. The colleague who also visited the church was there on both days to help and organize participants on the van. The first participation day was agreed to start at 7 pm because the preacher mentioned it was the best time for everyone. However, a week later he informed us that 5 pm would work just as well. It was easier to reserve the technology room for an earlier time, so this was the preferred option.

On September 10th, participants arrived 30 minutes later than expected at UFSC. We welcomed them and guided them to the technology room, which was in Block A, but had to be accessed through Block B due to security protocol at night. Each participant was seated in front of a computer. Firstly, we read and summed up the contents of the TCLE for everyone; only then the written lexical decision task was started. At first, task instructions would be given slowly to the entire group. However, this strategy did not work out because some people could not understand the instructions as well as the others. At this moment we found out that some participants had never used a computer before, so we had to teach some maneuvering first. Then, task instructions were given individually to each participant as they completed the tasks. Some participants got a piece of paper and a pen out of their pocket and started taking notes of the words which appeared during the task. The researcher never ordered this action and never considered it could happen. Then, I explicitly instructed participants not to take notes, to just focus on the task on the computer. Later three participants started making comments aloud on the tasks in HC with each other. Then, we asked them to refrain from commenting during the task and illustrated that it was just like a school exam. In addition, a participant took her toddler with her to the university since she had no one to leave her with. Moreover, accidentally two participants ended up turning off their computers directly from the wall plug during the written lexical decision task. At that moment, we decided it would be best for them to restart the task, since stimuli were presented randomly, and so they did. All these moments of distraction were registered in the data collection file so that data from the distracted participants could be removed from analyses. Due to all these delays, many participants were not able to finish the vocabulary test and to answer the language history questionnaire and the reading habits questionnaire. Data collection ended 30 minutes later than the scheduled time on September 10h. We informed the van driver and the guard that we would need some more time to wrap everything up, take everyone out of the building and tidy up the room. The guard was polite, but he explained that he could not allow any more delays from the original scheduled time. We agreed and every participant was taken back to the church by van.

On September 24th, participants also arrived 30 minutes later than expected. However, this time we had plenty of time scheduled. Only half of the expected number of people (20) were there; some did not show up, some arrived at the church later than the pick-up time. It was much easier to organize, help and manage the smaller group. One participant from the previous session — the one who had previously offered himself to invite people from his church — was there to help to communicate instructions and to solve doubts other participants might have during the tasks. This made data collection much smoother, and everyone was able to finish all tasks and questionnaires in time. Participants were very chatty and happy at the end and they thanked us a lot for the opportunity. All participants were offered cookies, juice and water at the end of the data collection session, outside of the computer room to compensate for the long duration of the session. We thanked them as well and let them know we would be available to talk about their participation at any time. Figure 66 illustrates how data collection was conducted.

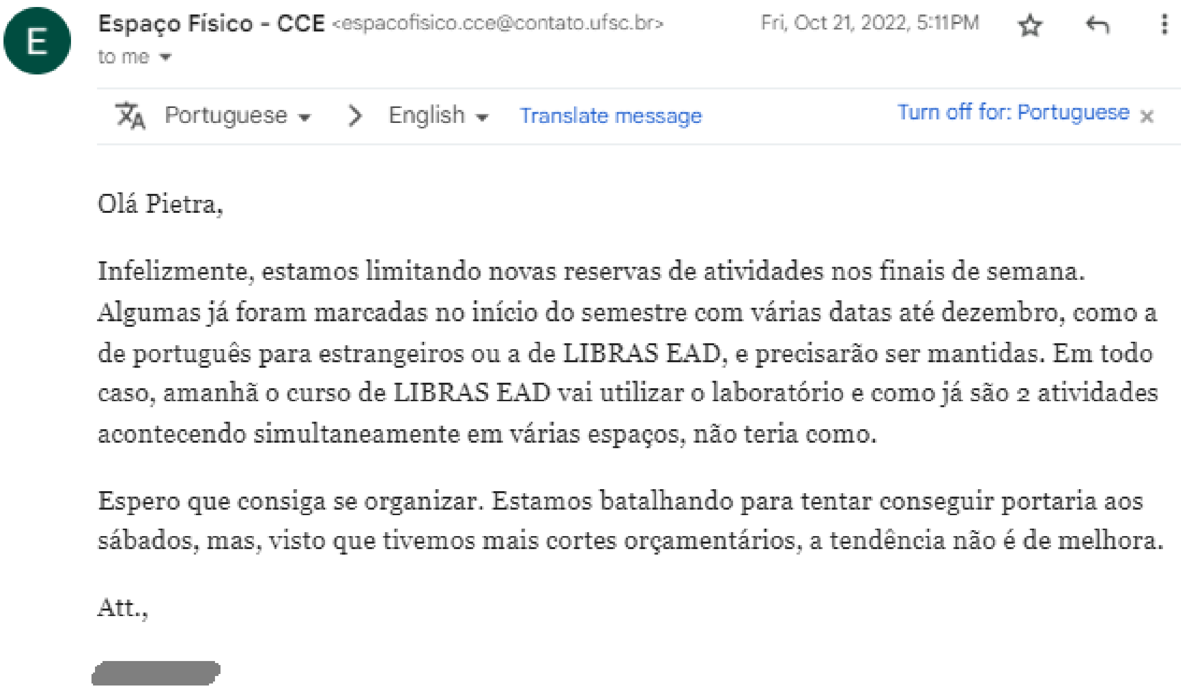
Figure 66 – Participant in Study 2 answering the questionnaire



Source: The Author (2024)

The plan was to have a third participation day for the ones who could not make it to the one on September 24th. However, the staff in charge of room reservation explained that it was not possible to have any new reservations on weekends other than the ones which were previously scheduled at the beginning of the semester. This decision was due to lack of staff to be available at weekends, as illustrated in Figure 67, an e-mail exchange with the head of room reservation. Another fact which impeded the continuation of data collection at UFSC was the sudden retraction in government funding for universities. Figures 68 and 69 are screenshots from two pieces of news reporting this. Data collection was planned to end in the middle of December 2022, which is also when usually school and university calendars end for the holidays. However, because of the impossibility of rescheduling the technology room at UFSC, it was canceled.

Figure 67 – E-mail exchange with room reservation staff



Espaço Físico - CCE <espacofisico.cce@contato.ufsc.br> Fri, Oct 21, 2022, 5:11PM

to me

Portuguese > English Translate message Turn off for: Portuguese

Olá Pietra,

Infelizmente, estamos limitando novas reservas de atividades nos finais de semana. Algumas já foram marcadas no início do semestre com várias datas até dezembro, como a de português para estrangeiros ou a de LIBRAS EAD, e precisarão ser mantidas. Em todo caso, amanhã o curso de LIBRAS EAD vai utilizar o laboratório e como já são 2 atividades acontecendo simultaneamente em várias espaços, não teria como.

Espero que consiga se organizar. Estamos batalhando para tentar conseguir portaria aos sábados, mas, visto que tivemos mais cortes orçamentários, a tendência não é de melhora.

Att.,

[redacted]

Source: The Author (2024)

Figure 68 – Piece of news from October 2022 about government investment retraction from universities

g1 ECONOMIA

DE OLHO NO ORÇAMENTO

Após bloqueio no MEC, reitores de universidades dizem que terão de 'cortar no osso': limpeza, restaurante, luz, água e bolsas estudantis

Ministério da Educação acumula bloqueio de quase R$ 3 bilhões em 2022 e é o ministério mais atingido pelos congelamentos, segundo dados da Instituição Fiscal Independente; R$ 763 milhões foram bloqueados nas universidades federais, o equivalente a quase 14% da verba anual das instituições, aponta Andifes.

Por Bianca Lima e Yasmim Perna, GloboNews — Brasília

06/10/2022 07h49 · Atualizado há 6 meses

Source: Lima and Perna (2022)

Figure 69 – Piece of news from November 2022 about government investment retraction from universities

uol INGRESSO.COM BATE-PAPO MEU NEGÓCIO PASSEI DIRETO PAGSEGURO UOL PLAY SAC EMAIL ENTRE ASSINE UOL

PRODUTOS NOTÍCIAS CARROS ECONOMIA FOLHA ESPORTE SPLASH UNIVERSA VIVABEM TILT ECOA CANAL UOL MOV NOSSA TAB UOL PRIME

REPORTAGEM

Governo faz corte de R$ 1,7 bi na verba do MEC em meio a jogo da seleção

Victor Godoy Veiga, ministro da Educação

Imagem: Luis Fortes/MEC

Chico Alves

Colunista do UOL

28/11/2022 21h48 | Atualizada em 28/11/2022 22h19

Source: Alves (2022)

On September 10th, 12 participants were not able to finish all tasks due to time restraints. Thus, I requested an amendment to the Ethics Committee so that I could adapt the questionnaires to a Google form version and could send it via WhatsApp to the participants who did not complete these tasks. Since it would not be possible to bring them back to the technology room at UFSC, and since most of them had smartphones, this would be the best option for them to answer the questionnaires. However, I could not get in contact with 8 participants; they never answered WhatsApp messages nor phone calls. In these cases, I insisted on trying to get information at least about their age and year of arrival in Brazil. I even asked the preacher to talk to them and see if they would answer these last questions; despite that, they never responded. When the data could not be gathered, this lack of information was replaced with the median for all participants in Study 2 when possible.

Other than the two sessions of data collection at UFSC with participants invited at a church, I visited two BP classes for foreigners at UFSC. These classes are free and happen twice a week. There were speakers of varied languages, such as Spanish, Arabic, and Haitian Creole. The teachers allowed and helped me to explain what my invitation was about, and I took notes of the e-mail address or the phone number of those interested in participating. There were 6 adult speakers of HC interested. I scheduled data collection sessions with them which would start at the end of the next BP class because it would be more convenient for them to meet at UFSC after their class. One of these 6 had only 1 hour available for participating due to work schedule and could only finish the lexical decision task and one of the sentence comprehension tasks. This person never answered whether they had interest in rescheduling the participation in order to finish all tasks, so I did not pursue their time further. Another of these 6 attended the scheduled participation session, but refused to participate after reading the TCLE. In the end, 4 of them participated.

Another 10 HC adult speakers were invited via acquaintances. For example, a friend or colleague a native HC speaker or a friend of a friend did, or a participant extended the participation invitation to a friend of theirs. This way, 4 participants completed the tasks remotely while asking me questions via WhatsApp or video call. Another 6 participants were able to meet in-person and completed tasks on my personal laptop. In this case, data collection was conducted in the quietest place available which would also be convenient for participants. For example, a small and quiet café near their workplace, a public library, or their home. In total, 44 adults completed tasks; however, I was able to analyze data from 35 of them because some participants displayed low attentiveness during data collection.

7 FINAL REMARKS

The general objective of this dissertation is to investigate the effects of cross-linguistic interaction between HC and BP on lexical access during comprehension of spoken and written BP by native speakers of HC. Two studies were designed; Study 1 had children and teenagers who spoke HC and BP as participants, while Study 2 had adults who spoke HC and BP as participants.

The specific objectives for Study 1 were:

1. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during visual and auditory lexical decision tasks in BP;
2. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce higher accuracy rates during visual and auditory lexical decision tasks in BP;
3. To investigate whether higher PA scores in BP, in comparison with lower PA scores, produce shorter RTs in visual and auditory lexical decision tasks in BP;
4. To investigate whether higher PA scores in BP, in comparison with lower PA scores, produce higher accuracy rates in visual and auditory lexical decision tasks in BP.

Hypotheses for Study 1 were:

1. Cognate words, in comparison with noncognate words, produce shorter RTs during both visual and auditory lexical decision tasks in BP.
2. Cognate words, in comparison with noncognate words, produce higher accuracy rates during both visual and auditory lexical decision tasks in BP.
3. Higher scores in PA in BP, in comparison with lower PA scores, produce shorter RTs in both visual and auditory lexical decision tasks in BP.
4. Higher scores in PA in BP, in comparison with lower PA scores, produce higher accuracy scores in both visual and auditory lexical decision tasks in BP.

In order to test these hypotheses, participants completed visual and auditory lexical decision tasks in BP and a PA test in BP. They also answered questions from the LHQ 3.0 and took a HC vocabulary test. Results showed that cognate words had no significant effect either on RTs nor on accuracy scores in any of the lexical decision tasks. So, Hypotheses 1 and 2 were refuted. However, PA scores were significantly and directly associated with accuracy scores, but not RTs. So, Hypothesis 3 was also refuted while Hypothesis 4 was indeed confirmed.

The specific objectives for Study 2 were:

1. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during a visual lexical decision task in BP;
2. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce higher accuracy rates during a visual lexical decision task in BP;
3. To investigate whether cognate words across HC and BP, in comparison with noncognate words, produce shorter RTs during visual and auditory sentence comprehension tasks in BP;
4. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during a visual lexical decision task in BP;
5. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce higher accuracy rates during a visual lexical decision task in BP;
6. To investigate whether more frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during visual and auditory sentence comprehension tasks in BP.

Hypotheses for Study 2 were:

1. Cognate words, in comparison with noncognate words, produce shorter RTs during a visual lexical decision task in BP;
2. Cognate words, in comparison with noncognate words, produce higher accuracy rates during a visual lexical decision task in BP;
3. Cognate words, in comparison with noncognate words, produce shorter RTs during both visual and auditory sentence comprehension tasks in BP;
4. More frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during a visual lexical decision task in BP;
5. More frequent RH in BP, in comparison with less frequent RH, produce higher accuracy rates during a visual lexical decision task in BP;

6. More frequent RH in BP, in comparison with less frequent RH, produce shorter RTs during both visual and auditory sentence comprehension tasks in BP.

In order to test these hypotheses, participants completed the same visual lexical decision task from study 1, visual and auditory self-paced sentence comprehension tasks in BP and a RH in BP questionnaire. They also answered questions from the LHQ 3.0 and took a HC vocabulary test. Results showed that cognate words had no significant effect either on RTs nor on accuracy scores in the visual lexical decision task. They also seemed to not affect RTs in either of the self-paced sentence comprehension tasks. So, hypotheses 1 and 2 were refuted. In addition, RH in BP were not associated with any measures in any of the tasks. Hypotheses 3 and 4 were refuted as well.

In sum, cognate words across HC and BP did not have any effect during language processing by immigrant bilingual children, teenagers, and adults. This result could offer evidence to a selective bilingual language processing account. However, because of limitations of the study and of my proposed interpretation involving lack of lexical knowledge, which will be repeated below, I understand this study does not present evidence to either the selective or the non-selective account.

There were two important limitations in terms of stimuli selection for my studies. The first one was frequency of occurrence and the second one was length. Due to difficulties in finding an annotated and normalized HC corpus, I constructed a compilation of HC words and their translations in BP, with information about parts of speech, cognate status across languages and relative frequency according to a highly educated speaker of HC who has lived in Brazil since 2014. Despite efforts to control for word frequency, this is still a subjective measure based on one speaker's perception. In addition, I could not restrict the length of words due to the number of cognate ones available. So, after adding word length to statistical models, it did not seem to have an effect in RTs. Longer words presented longer RTs, which was expected. However, no interaction between length and cognate status was seen. Thus, word length did not seem to have influenced the absence of a cognate effect.

Moreover, there were two other limitations which appeared during data collection. In Study 1, participation sessions were carried out in any room which received the least amount of noise in the schools. This means that in many cases it was not possible to have a quiet space for participants to perform the tasks. Headphones and earphones were always available, but sometimes noise was unavoidable. In Study 2, during data collection many participants demonstrated not knowing words in the sentence tasks, including the comprehension questions.

Sometimes they stopped in the middle of the task to ask the researcher about the meaning of a word. The researcher would explain that she could not tell them the meaning at that moment and that they should try to complete the task with information from the whole sentence. So, Icould not fully trust their answers for the comprehension questions. Despite that, I included an exclusion criterion for accuracy on the comprehension questions. Also, I did take notes about moments of distraction during data collection and removed data from specific participants.

After considering these limitations, I discussed the results in relation to relative language activation and to lexical knowledge. The parallel language activation seems to be modulated by proficiency among other factors (Dijkstra, 2005). The language co-activation effects appear between a minimum level of proficiency and a maximum one in the L2. So, beginners and highly advanced speakers may not show cognate effects for example. In both studies, I saw no language co-activation effect despite the fact that participants reported intermediate levels of proficiency in BP. Thus, I turned to my tentative explanation of the absence of effects.

My speculation was based on two theoretical accounts and on evidence from studies. I presented the learning account (Costa *et al.*, 2017), which posits that the structure and organization of the L1 guides the learning of the L2. When a bilingual person is learning a word such as a cognate one, orthographic, phonological and semantic representations from both languages become connected because of the similarities between the word from the L2 and the one from the L1. Then, I linked this account to the Lexical Quality Hypothesis (Perfetti; Hart, 2001), which states when a person has words with high-quality lexical representations, reading becomes more efficient. A word is of high quality when its constituents — orthography, phonology, grammar, and meaning — are stable, fully-specified and well-bound. This high standard can be achieved through a rich and varied experience with words. Finally, I directed the argument to evidence showing that SES levels have an impact on language development (Corso *et al.* 2016). This reasoning was an attempt to explain both the absence of language co-activation effects and the variation in accuracy scores. However, my study design does not offer any possibility to test these ideas, and they should be considered cautiously. The debate presented in this work should be taken as inspiration for future studies instead of solid claims about bilingual language processing.

In relation to next steps in this topic of research, I observed at least three aspects which could be detailed and expanded. Firstly, in this work, more studies could have been used towards the construction of a comprehensive, annotated and normalized corpus of HC. I did find materials which could be merged into the organization of a corpus. However, they were individual files separated according to source and year. Thus, a compiled version containing all that information would be easier to use and more reliable. Secondly, there is still the need to keep testing the limits of language co-activation and of the cognate effect in minority languages. Many studies evaluate the development of language skills for minority language children (for example, Verhoeven, 2007), but not as many consider the literature on language co-activation for children and adults in the context of minority languages. The prevalence of bottom-up effects in bilingual lexical processing which is predicted by most of the literature was not seen here and was observed only in part (Woolpert, 2018). Finally, based on my tentative interpretation of results, I point out that studies about language co-activation could have more diverse samples, for example participants from different SES backgrounds. This way, the bilingual word identification system can be tested in relation to varied circumstances of language input and print exposure. More specifically, future studies could consider characteristics of bilingual participants who have diverse levels of richness in word experience (Taylor; Perfetti, 2016) from different SES backgrounds and verify whether parallel language activation is indeed modulated by these external factors.

The research process which culminated in this study report is in consonance with statements about the complexity intrinsic to bilingualism and multilingualism. The studies presented here were not able to disentangle most relationships as intended in the objectives. However, I did observe the influence of L2 phonological awareness during L2 lexical access as shown in the literature. I also contribute with suggestions to future studies, which are plausible when bilingualism is considered as a continuum of diverse experiences instead of one homogenous category (Poarch; Bialystok, 2015) and when social and environmental factors are considered in language learning and development (Titone; Tiv, 2023). Complexity demands the inclusion of less investigated variables and more innovative and replication studies. And there are still plenty of languages and linguistic phenomena in the world to investigate.

REFERENCES

ABC – AGÊNCIA BRASILEIRA DE COOPERAÇÃO. Haiti. **Agência Brasileira de Cooperação**: Ministério das Relações Exteriores, Brasília, 2012. Available at: http://www.abc.gov.br/projetos/cooperacaosulsul/haiti. Last access on: 13 Mar 2024.

ACHESON, D. J.; GANUSHCHAK, L. Y.; CHRISTOFFELS, I. K.; HAGOORT, P. Conflict monitoring in speech production: physiological evidence from bilingual picture naming. **Brain and Language**, [*S. l.*], v. 123, n. 2, p. 131-136, 2012. DOI: https://doi.org/10.1016/j.bandl.2012.08.008. Last access on: 13 Mar 2024.

ACNUR – ALTO COMISSARIADO DAS NAÇÕES UNIDAS PARA REFUGIADOS; CSVM – CÁTEDRA SÉRGIO VIEIRA DE MELLO. **Relatório Anual 2022**. Brasília: ACNUR; CSVM, 2002. Available at: https://www.acnur.org/portugues/wp-content/uploads/2022/09/CSVM-Relato%CC%81rio-Anual-2022.pdf. Last access on: 13 Mar 2024.

ADIS-CASTRO, G.; POSTMAN, L. Psychophysical methods in the study of word recognition. **Science**, [*S. l.*], v. 125, n. 3240, p. 193-194, 1957. DOI: https://doi.org/10.1126/science.125.3240.193. Last access on: 13 Mar 2024.

ALLEN, D. Cognate frequency and assessment of second language lexical knowledge. **International Journal of Bilingualism**, [*S. l.*], v. 23, n. 5, p. 1121-1136, 2019. DOI: https://doi.org/10.1177/1367006918781063. Last access on: 13 Mar 2024.

ALLEN, D.; CONKLIN, K.; MIWA, K. Cross-linguistic lexical effects in different-script bilingual reading are modulated by task. **International Journal of Bilingualism**, [*S. l.*], v. 25, n. 1, p. 168-188, 2021. DOI: https://doi.org/10.1177/1367006920943974. Last access on: 13 Mar 2024.

ALTENHOFEN, C. V. Bases para uma política linguística das línguas minoritárias no Brasil. In: NICOLAIDES, C.; SILVA, K. A.; TÍLIO, R.; ROCHA, H. C. (org.). **Política e políticas linguísticas**. Campinas: Pontes Editores, 2013. p. 93-116.

ALVES, C. Governo faz corte de R$ 1,7 bi na verba do MEC em meio a jogo da seleção. **UOL**, [*S. l.*], 28 Nov 2022. Available at: https://noticias.uol.com.br/colunas/chico-alves/2022/11/28/governo-faz-corte-de-r-17-bi-na-verba-do-mec-em-meio-a-jogo-da-selecao.htm. Last access on: 13 Mar 2024.

ALVES, I. S.; OLIVEIRA, J. B. A.; HIRATA, G. **Perdas na pandemia**: alfabetização, leitura e fluência de leitura – dados de Sobral. Rio de Janeiro: Instituto Alfa e Beto; Idados, 2022. Available at: https://www.alfaebeto.org.br/wp-content/uploads/2022/05/Nota-Tecnica.pdf. Last access on: 13 Mar 2024.

AMORIM, A. N.; JEON, L.; ABEL, Y.; FELISBERTO, E. F.; BARBOSA, L. N. F.; DIAS, N. M. Using Escribo Play video games to improve phonological awareness, early reading, and writing in preschool. **Educational Researcher**, [*S. l.*], v. 49, n. 3, p. 188-197, 2020. DOI: https://doi.org/10.3102/0013189X20909824. Last access on: 13 Mar 2024.

ANDREWS, S. Frequency and neighborhood effects on lexical access: activation or search? **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 15, n. 5, p. 802-814, 1989. DOI: https://doi.org/10.1037/0278-7393.15.5.802. Last access on: 13 Mar 2024.

ANDREWS, S. Frequency and neighborhood effects on lexical access: lexical similarity or orthographic redundancy? **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 18, n. 2, p. 234-254, 1992.

ANGERMEYER, P. S. Interpreter-mediated interaction as bilingual speech: bridging macro- and micro-sociolinguistics in codeswitching research. **International Journal of Bilingualism**, [*S. l.*], v. 14, n. 4, p. 466-489, 2010. DOI: https://doi.org/10.1177/1367006910370914. Last access on: 13 Mar 2024.

ANTHONY, J. L.; FRANCIS, D. J. Development of phonological awareness. **Current Directions in Psychological Science**, [*S. l.*], v. 14, n. 5, p. 255-259, 2005. DOI: https://doi.org/10.1111/j.0963-7214.2005.00376.x. Last access on: 13 Mar 2024.

ANTHONY, J. L.; LONIGAN, C. J. The nature of phonological awareness: converging evidence from four studies of preschool and early grade school children. **Journal of Educational Psychology**, [*S. l.*], v. 96, n. 1, p. 43-55, 2004. DOI: https://doi.org/10.1037/0022-0663.96.1.43. Last access on: 27 May 2024.

ANTHONY, J. L.; LONIGAN, C. J.; DRISCOLL, K.; PHILLIPS, B. M.; BURGESS, S. R. Phonological sensitivity: a quasi-parallel progression of word structure units and cognitive operations. **Reading Research Quarterly**, [*S. l.*], v. 38, n. 4, p. 470-487, 2003. DOI: https://doi.org/10.1598/RRQ.38.4.3. Last access: 27 May 2024.

ANTÓN, E.; DUÑABEITIA, J. A. Better to be alone than in bad company: cognate synonyms impair word learning. **Behavioral Sciences**, [*S. l.*], v. 10, n. 8, p. 1-23, 2020. DOI: https://doi.org/10.3390/bs10080123. Last access on: 13 Mar 2024.

APA – AMERICAN PSYCHOLOGICAL ASSOCIATION. Education and socioeconomic status, resources and publications. **American Psychological Association**, [*S. l.*], 2017. Available at: https://www.apa.org/pi/ses/resources/publications/education. Last access on: 13 Mar 2024.

ARAÚJO, S. G.; MARTINS-REIS, V. de O.; COUTO, E. de A. B.; ALVES, L. M. Desempenho de adolescentes no discurso narrativo oral e fatores associados. **CoDAS**, [*S. l.*], v. 34, n. 5, e20210207, 2022. DOI: https://doi.org/10.1590/2317-1782/20212021207. Last access on: 13 Mar 2024.

ARCHER, J.; CHAMPION, T.; TYRONE, M. E.; WALTERS, S. Phonological development of monolingual Haitian Creole-speaking preschool children. **Communication Disorders Quarterly**, [*S. l.*], v. 39, n. 3, p. 426-437, 2018.

ARÊAS DA LUZ FONTES, A. B.; BRENTANO, L. de S.; TOASSI, P. F. P.; SITTIG, C. C.; FINGER, I. Evidence of non-selective lexical access in children from a Portuguese-English bilingual school. **Prolíngua**, João Pessoa, v. 15, n. 2, p. 183-197, 2020. DOI: https://doi.org/10.22478/ufpb.1983-9979.2020v15n2.54901. Last access on: 13 Mar 2024.

ARÊAS DA LUZ FONTES, A. B.; YEH, L.-H.; SCHWARTZ, A. I. Desambiguação lexical bilíngue: a natureza dos efeitos de coativação lexical entre as línguas. **Letrônica**, Porto Alegre, v. 3, n. 1, p. 107-127, 2010. Available at: https://revistaseletronicas.pucrs.br/index.php/letronica/article/view/7074. Last access on: 13 Mar 2024.

ARÊAS DA LUZ FONTES; A. B.; SCHWARTZ, A. I. Bilingual access of homonym meanings: individual differences in bilingual access of homonym meanings. **Bilingualism**: Language and Cognition, [*S. l.*], v. 18, n. 4, p. 639-656, 2015. DOI: https://doi.org/10.1017/S1366728914000509. Last access on: 13 Mar 2024.

ARTIEDA, G. The role of L1 literacy and reading habits on the L2 achievement of adult learners of English as a foreign language. **System**, [*S. l.*], n. 66, p. 168-176, 2017.

BAENINGER, R. Imigração haitiana no Brasil. *In*: BAENINGER, R.; PERES, R.; FERNANDES, D.; DA SILVA, S. A.; ASSIS, G. O.; CASTRO, M. C. G.; COTINGUIBA, M. P. (org.). **Imigração haitiana no Brasil**. Jundiaí: Paco e Littera, 2017. Chapter 27. Available at: https://vdoc.com.br/book/2rwwwzqj09rv/imigraao-haitiana-no-brasil. Last access on: 13 Mar 2024.

BARBOSA, I. V.; FISTAROL, C. F. S.; SILVEIRA, É. L. Línguas de herança e de imigração: notas sobre comunidades linguísticas e (co)existência sociocultural no Brasil. **Água Viva**, [*S. l.*], v. 5, n. 3, s. p., 2020. DOI: 10.26512/aguaviva.v5i3.35812. Disponível em: https://periodicos.unb.br/index.php/aguaviva/article/view/35812. Last access: 25 May 2024.

BARTON, J. J.; HANIF, H. M.; EKLINDER BJÖRNSTRÖM, L.; HILLS, C. The word-length effect in reading: a review. **Cognitive neuropsychology**, [*S. l.*], v. 31, n. 5-6, p. 378-412, 2014. DOI: https://doi.org/10.1080/02643294.2014.895314. Last access on: 13 Mar 2024.

BATES, D.; MÄCHLER, M.; BOLKER, B.; WALKER, S. fitting linear mixed-effects models using lme4. **Journal of Statistical Software**, [*S. l.*], v. 67, n. 1, p. 1-48, 2015. DOI: https://doi.org/10.18637/jss.v067.i01. Last access on: 13 Mar 2024.

BEATTY-MARTÍNEZ, A. L.; NAVARRO-TORRES, C. A.; DUSSIAS, P. E.; BAJO, M. T.; GUZZARDO TAMARGO, R. E.; KROLL, J. F. Interactional context mediates the consequences of bilingualism for language and cognition. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 46, n. 6, p. 1022-1047, 2020. DOI: https://doi.org/10.1037/xlm0000770.

BLOOMFIELD, L. **Language**. New York: Holt, 1933.

BOHN, O.-S. Cross-language and second language speech perception. *In*: FERNÁNDEZ, E. M.; CAIRNS, H. S. (ed.). **The handbook of psycholinguistics**. Hoboken: Wiley Blackwell, 2018. p. 213-239. DOI: https://doi.org/10.1002/9781118829516.ch10. Last access on: 13 Mar 2024.

BONCOURT, T.; RAVECCA, P. Power, politics, and the development of political science in the Americas. **Science in Context**, [*S. l.*], v. 33, n. 2, p. 95-100, 2020. DOI: https://doi.org/10.1017/S0269889720000150. Last access on: 13 Mar 2024.

BOSMA, E.; BLOM, E.; HOEKSTRA, E.; VERSLOOT, A. A longitudinal study on the gradual cognate facilitation effect in bilingual children's Frisian receptive vocabulary. **International Journal of Bilingual Education and Bilingualism**, [*S. l.*], v. 22, n. 4, p. 371-385, 2019. DOI: https://doi.org/10.1080/13670050.2016.1254152. Last access on: 13 Mar 2024.

BOSMA, E.; NOTA, N. Cognate facilitation in Frisian-Dutch bilingual children's sentence reading: an eye-tracking study. **Journal of Experimental Child Psychology**, [*S. l.*], v. 189, e104699, 2020. DOI: https://doi.org/10.1016/j.jecp.2019.104699. Last access on: 13 Mar 2024.

BOUKRINA, O.; HANSON, S. J.; HANSON, C. Modeling activation and effective connectivity of VWFA in same script bilinguals. **Human Brain Mapping**, [*S. l.*], v. 35, n. 6, p. 2543-2560, 2014. DOI: https://doi.org/10.1002/hbm.22348. Last access on: 13 Mar 2024.

BOULTON, G. S. **Science as a global public good**. Paris: International Science Council Position Paper, 2021. Available at: https://council.science/wp-content/uploads/2020/06/Science-as-a-global-public-good_v041021.pdf. Last access on: 13 Mar 2024.

BRANZI, F. M.; MARTIN, C. D.; BIAU, E. Activating words without language: beta and theta oscillations reflect lexical access and control processes during verbal and non-verbal object recognition tasks. **Cerebral Cortex**, [*S. l.*], v. 33, n. 10, p. 6228-6240, 2023. DOI: https://doi.org/10.1093/cercor/bhac499. Last access on: 13 Mar 2024.

BRANZI, F. M.; MARTIN, C. D.; CARREIRAS, M.; PAZ-ALONSO, P. M. Functional connectivity reveals dissociable ventrolateral prefrontal mechanisms for the control of multilingual word retrieval. **Human Brain Mapping**, [*S. l.*], v. 41, n. 1, p. 80-94, 2020. DOI: https://doi.org/10.1002/hbm.24788. Last access on: 13 Mar 2024.

BRASIL. Português como língua de herança. **Ministério das Relações Exteriores**: Consulado-Geral do Brasil em Londres, [*S. l.*], 2013. Available at: https://www.gov.br/mre/pt-br/consulado-londres/portugues-como-lingua-de-heranca-1#:~:text=O%20conceito%20L%C3%ADngua%20de%20Heran%C3%A7a,s)%20e%20aos%20valores%20familiares. Last access on: 13 Mar 2024.

BRAZE, D.; GONG, T. Orthography, word recognition, and reading. *In*: FERNANDEZ, E. M.; CAIRNS, H. S. (ed.). **Handbook of Psycholinguistics**. Hoboken: Wiley-Blackwell, 2018. p. 269-293.

BREHM, L.; ALDAY, P. M. Contrast coding choices in a decade of mixed models. **Journal of Memory and Language**, [*S. l.*], v. 125, e104334, 2022. DOI: https://doi.org/10.1016/j.jml.2022.104334. Last access on: 13 Mar 2024.

BREITER, W. **Rapid bootstrapping of Haitian Creole large vocabulary continuous speech recognition**. Karlsruher: Karlsruher Institut für Technologie, 2014. Available at:

https://www.csl.uni-bremen.de/cms/publications/bibtexbrowser.php?key=breiter2014rapid&bib=csl_all_publications.bib. Last access on: 13 Mar 2024.

BRENDERS, P.; VAN HELL, J. G.; DIJKSTRA, T. Word recognition in child second language learners: evidence from cognates and false friends. **Journal of Experimental Child Psychology**, [*S. l.*], v. 109, n. 4, p. 383-396, 2011. DOI: https://doi.org/10.1016/j.jecp.2011.03.012. Last access on: 13 Mar 2024.

BRIL, M.; GERRITS, A.; VISSER, M. The effects of linguistic and cognitive factors on the L2 processing of oral input: a self-paced listening experiment. **International Journal of Listening**, [*S. l.*], v. 36, n. 3, p. 268-282, 2022. DOI: https://www.doi.org/10.1080/10904018.2021.1992281. Last access on: 13 Mar 2024.

BROERSMA, M.; CARTER, D.; ACHESON, D. J. Cognate costs in bilingual speech production: evidence from language switching. **Frontiers in Psychology**, [*S. l.*], v. 27, p. 1-16, 2016. DOI: https://doi.org/10.3389/fpsyg.2016.01461. Last access on: 13 Mar 2024.

BRYSBAERT, M.; STEVENS, M. Power analysis and effect size in mixed effects models: a tutorial. **Journal of Cognition**, [*S. l.*], v. 1, n. 1, p. 1-20, 2018. DOI: https://doi.org/10.5334/joc.10.

BUCHWEITZ, A.; MOTA, M.; NAME, C. Linguagem: das primeiras palavras à aprendizagem da leitura. *In*: LENT, R.; BUCHWEITZ; A.; MOTA, M. (org.). **Ciência para educação**: uma ponte entre dois mundos. São Paulo: Atheneu, 2018. v. 1, p. 119-132.

BULTENA, S.; DIJKSTRA, T.; VAN HELL, J. G. Cognate effects in sentence context depend on word class, L2 proficiency, and task. **Quarterly Journal of Experimental Psychology**, [*S. l.*], v. 67, n. 6, p. 1214-1241, 2014. DOI: https://doi.org/10.1080/17470218.2013.853090. Last access on: 13 Mar 2024.

BULTENA, S.; DIJKSTRA, T.; VAN HELL, J. G. Switch cost modulations in bilingual sentence processing: evidence from shadowing. **Language, Cognition and Neuroscience**, [*S. l.*], v. 30, n. 5, p. 586-605, 2015. DOI: https://doi.org/10.1080/23273798.2014.964268. Last access on: 13 Mar 2024.

BUTIKOFER, E. A.; SILVA, E. A. da. Imigração e periferias urbanas: experiências haitianas em São Paulo. **REMHU**: Revista Interdisciplinar da Mobilidade Humana, Brasília, v. 29, n. 62, p. 151-169, 2021. DOI: https://doi.org/10.1590/1980-85852503880006210. Last access on: 13 Mar 2024.

BUTLER, Y. G. Kanji Acquisition among language minority students in Japan: a comparative study of Japanese-as-a-second-language students born in Japan. **Working Papers in Educational Linguistics**, [*S. l.*], v. 26, n. 1, p. 1-20, 2011. Available at: https://wpel.gse.upenn.edu/sites/default/files/Butler.pdf. Last access on: 13 Mar 2024.

BYERS-HEINLEIN, K.; LEW-WILLIAMS, C. Language comprehension in monolingual and bilingual children. *In*: FERNÁNDEZ, E. M.; CAIRNS, H. S. (ed.). **The handbook of Psycholinguistics**. Hoboken: Wiley Blackwell, 2018. p. 516-535. DOI: https://doi.org/10.1002/9781118829516.ch23. Last access on: 13 Mar 2024.

BYRNE, B.; FIELDING-BARNSLEY, R. Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. **Journal of Educational Psychology**, [*S. l.*], v. 81, n. 3, p. 313-321, 1989. DOI: https://doi.org/10.1037/0022-0663.81.3.313. Last access on: 13 Mar 2024.

CAI, Z. G.; PICKERING, M. J.; YAN, H.; BRANIGAN, H. P. Lexical and syntactic representations in closely related languages: evidence from Cantonese-Mandarin bilinguals. **Journal of Memory and Language**, [*S. l.*], v. 65, n. 4, p. 431-445, 2011. DOI: https://doi.org/10.1016/j.jml.2011.05.003. Last access on: 13 Mar 2024.

CAMICIOTTOLI, B. C. Extensive reading in English: habits and attitudes of a group of Italian university EFL students. **Journal of Research in Reading**, [*S. l.*], v. 24, n. 2, p. 135-153, 2002. DOI: https://doi.org/10.1111/1467-9817.00137. Last access on: 13 Mar 2024.

CAPELLINI, S. A.; OLIVEIRA, A. M.; CUETOS, F. **PROLEC**: Provas de Avaliação dos Processos de Leitura – Manual. 2. ed. São Paulo, SP: Casa do Psicólogo, 2012.

CAPES – COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. Programa Emergencial Pró-Haiti. **Ministério da Educação**, Brasília, 6 Jun 2011. Available at: https://www.gov.br/capes/pt-br/acesso-a-informacao/acoes-e-programas/bolsas/bolsas-e-auxilios-internacionais/informacoes-internacionais/programas-encerrados-internacionais/programa-emergencial-pro-haiti. Last access on: 13 Mar 2024.

CAPOVILLA, F. C.; SEABRA, A. G. Prova de consciência fonológica por escolha de figuras. *In*: SEABRA, A. G.; DIAS, N. M. (org.). **Avaliação neuropsicológica cognitiva**: linguagem oral. São Paulo: Memnon Edições Científicas, 2013. v. 2, chapter 16. (*E-book*).

CARAVOLAS, M.; BRUCK, M. The effect of oral and written language input on children's phonological awareness: a cross-linguistic study. **Journal of Experimental Child Psychology**, [*S. l.*], v. 55, n. 1, p. 1-30, 1993. DOI: https://doi.org/10.1006/jecp.1993.1001. Last access on: 13 Mar 2024.

CARROLL, J. M.; SNOWLING, M. J.; STEVENSON, J.; HULME, C. The development of phonological awareness in preschool children. **Developmental Psychology**, [*S. l.*], v. 39, n. 5, p. 913-923, 2003. DOI: https://doi.org/10.1037/0012-1649.39.5.913. Last access: 27 May 2024.

CASSOL RIGATTI, P.; ARÊAS DA LUZ FONTES, A. B. Cognate and meaning frequency effects on homonym processing. **International Journal of Bilingualism**, [*S. l.*], v. 26, n. 6, p. 732-748, 2022. DOI: https://doi.org/10.1177/13670069211060729. Last access on: 13 Mar 2024.

CASTLES, A.; RASTLE, K.; NATION, K. Ending the reading wars: reading acquisition from novice to expert. **Psychological Science in the Public Interest**, [*S. l.*], v. 19, p. 5-51, 2018.

CATTS, H. W.; HERRERA, S.; NIELSEN, D. C.; BRIDGES, M. S. Early prediction of reading comprehension within the simple view framework. **Reading and Writing**, [*S. l.*], v. 28, p. 1407-1425, 2015. DOI: https://doi.org/10.1007/s11145-015-9576-x. Last access on: 13 Mar 2024.

CAVALCANTI, L.; DE OLIVEIRA, T.; SILVA, B. G. (org.) **Resumo Executivo 2022**. [*S. l.*]: Observatório das Migrações Internacionais (OBMigra), 2022. Available at: https://portaldeimigracao.mj.gov.br/images/Obmigra_2020/OBMigra_2022/RELAT%C3%93RIO_ANUAL/Resumo_Executivo_2022_-_Versa%CC%83o_completa_01.pdf. Last access on: 13 Mar 2024.

CAVICCHIOLO, E.; MANGANELLI, S.; BIANCHI, D.; BIASI, V.; LUCIDI, L.; GIRELLI, L.; COZZOLINO, M.; ALIVERNINI, F. Social inclusion of immigrant children at school: the impact of group, family and individual characteristics, and the role of proficiency in the national language. **International Journal of Inclusive Education**, [*S. l.*], v. 27, n. 2, p. 146-166, 2023. DOI: https://doi.org/10.1080/13603116.2020.1831628. Last access on: 13 Mar 2024.

CAYLAK TOPLU, N.; ERTEN, I. H. A study into the interplay between first and second language reading motivation, reading habits and vocabulary size. **Sustainable Multilingualism [Darnioji daugiakalbystė]**, [*S. l.*], v. 22, p. 94-137, 2023. DOI: https://doi.org/10.2478/sm-2023-0005. Last access on: 13 Mar 2024.

CAZENEUVE, M.; NASCIMENTO, L. C. R. A influência da situação linguística do Haiti no processo de aprendizagem das crianças. **Revista Digital de Políticas Lingüísticas**, [*S. l.*], v. 8, p. 71-87, 2016.

CDC – CENTERS FOR DISEASE CONTROL AND PREVENTION. End of the federal COVID-19 public health emergency (PHE) declaration. **CDC Archive**, [*S. l.*], 12 Sep 2023. Available at: https://www.cdc.gov/coronavirus/2019-ncov/your-health/end-of-phe.html#:~:text=The%20federal%20COVID%2D19%20PHE,to%20align%20with%20data%20changes. Last access on: 13 Mar 2024.

CFN – CORPO DE FUZILEIROS NAVAIS. MINUSTAH. **Corpo de Fuzileiros Navais**: Marinha do Brasil, [*S. l.*], [201-?]. Available at: https://www.marinha.mil.br/cgcfn/minustah. Last access on: 13 Mar 2024.

CHAMBERS, C. G.; COOKE, H. Lexical competition during second-language listening: sentence context, but not proficiency, constrains interference from the native lexicon. **Journal of Experimental Psychology**: Learning, memory, and cognition, [*S. l.*], v. 35, n. 4, p. 1029-1040, 2009. DOI: https://doi.org/10.1037/a0015901. Last access on: 13 Mar 2024.

CHARLES, J. Dez anos após ser devastado por terremoto, Haiti ainda vive em cenário desolador. **National Geographic**: História, [*S. l.*], 15 Jan 2020. Available at: https://www.nationalgeographicbrasil.com/historia/2020/01/dez-anos-apos-ser-devastado-por-terremoto-haiti-ainda-vive-em-cenario-desolador. Last access on: 13 Mar 2024.

CHEN, S.; FU, T.; ZHAO, M.; ZHANG, Y.; PENG, Y.; YANG, L.; GU, X. Masked translation priming with concreteness of cross-script cognates in visual word recognition by Chinese learners of English: an ERP study. **Frontiers in Psychology**, [*S. l.*], v. 12, e796700, 2022. DOI: https://doi.org/10.3389/fpsyg.2021.796700. Last access on: 13 Mar 2024.

CHRISTENSEN, L. B.; JOHNSON, B.; TURNER, L. A. **Research methods, design, and analysis**. 12. ed. Pearson: Global Edition, 2015.

CHRISTOFFELS, I. K.; FIRK, C.; SCHILLER, N. O. Bilingual language control: an event-related brain potential study. **Brain Research**, [*S. l.*], v. 1147, p. 192-208, 2007. DOI: https://doi.org/10.1016/j.brainres.2007.01.137. Last access on: 13 Mar 2024.

CLYNE, M. Multilingualism. *In*: COULMAS, F. (ed.). **The handbook of Sociolinguistics**. Malden: Blackwell Publishing Ltd, 2017. Chapter 18. DOI: https://doi.org/10.1002/9781405166256.ch18. Last access on: 13 Mar 2024.

CNIG – CONSELHO NACIONAL DE IMIGRAÇÃO. Listagem de permanência – haitianos. **CNIg**, Brasília, 12 Nov 2015. Available at: https://portaldeimigracao.mj.gov.br/pt/listagem-de-permanencia-haitianos. Last access on: 13 Mar 2024.

COLTHEART, M.; RASTLE, K.; PERRY, C.; LANGDON, R.; ZIEGLER, J. DRC: a dual route cascaded model of visual word recognition and reading aloud. **Psychological review**, [*S. l.*], v. 108, n. 1, p. 204-256, 2001. DOI: https://doi.org/10.1037/0033-295x.108.1.204. Last access on: 13 Mar 2024.

COMESAÑA, M.; FERRÉ, P.; ROMERO, J.; GUASCH, M.; SOARES, A. P.; GARCÍA-CHICO, T. Facilitative effect of cognate words vanishes when reducing the orthographic overlap: the role of stimuli list composition. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 41, n. 3, p. 614-635, 2015.

COMESAÑA, M.; HARO, J.; MACIZO, P.; FERRÉ, P. Disentangling the role of deviant letter position on cognate word processing. **Frontiers in Psychology**, [*S. l.*], v. 12, e731312, 2021. DOI: https://doi.org/10.3389/fpsyg.2021.731312. Last access on: 13 Mar 2024.

COMESAÑA, M.; SÁNCHEZ-CASAS, R.; SOARES, A. P.; PINHEIRO, A. P.; RAUBER, A.; FRADE, S.; FRAGA, I. The interplay of phonology and orthography in visual cognate word recognition: an ERP study. **Neuroscience Letters**, [*S. l.*], v. 529, n. 1, p. 75-79, 2012. DOI: https://doi.org/10.1016/j.neulet.2012.09.010. Last access on: 13 Mar 2024.

CONSED – CONSELHO NACIONAL DE SECRETÁRIOS DE EDUCAÇÃO. Programa garante apoio pedagógico para alunos migrantes da rede estadual [Santa Catarina]. **CONSED**, Brasília, 11 May 2022. Available at: https://www.consed.org.br/noticia/programa-garante-apoio-pedagogico-para-alunos-migrantes-da-rede-estadual-de-educacao-de-sc. Last access on: 13 Mar 2024.

COOPER, D. H.; ROTH, F. P.; SPEECE, D. L.; SCHATSCHNEIDER, C. The contribution of oral language skills to the development of phonological awareness. **Applied Psycholinguistics**, [*S. l.*], v. 23, n. 3, p. 399-416, 2002. DOI: https://doi.org/10.1017/S0142716402003053. Last access on: 13 Mar 2024.

COOPER, F. S.; LIBERMAN, A. M.; BORST, J. M. The interconversion of audible and visible patterns as a basis for research in the perception of speech. **Proceedings of the National Academy of Science**, [*S. l.*], v. 37, p. 318-325, 1951.

COP, U.; DIRIX, N.; VAN ASSCHE, E.; DRIEGHE, D.; DUYCK, W. Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading.

Bilingualism: Language and Cognition, [*S. l.*], v. 20, n. 4, p. 1-23, 2016. DOI: https://doi.org/10.1017/S1366728916000213. Last access on: 13 Mar 2024.

CORSO, H. V.; CROMLEY, J. G.; SPERB, T.; SALLES, J. F. Modeling the relationship among reading comprehension, intelligence, socioeconomic status, and neuropsychological functions: the mediating role of executive functions. **Psychology & Neuroscience**, [*S. l.*], v. 9, n. 1, p. 32-45, 2016. DOI: https://doi.org/10.1037/pne0000036. Last access on: 13 Mar 2024.

CORTESE, M. J.; BALOTA, D. A. Visual word recognition in skilled adult readers. *In*: SPIVEY, M. J.; MCRAE, K.; JOANISSE, M. F. (ed.). **The Cambridge handbook of Psycholinguistics**. [*S. l.*]: Cambridge University Press, 2012. p. 159-185. DOI: https://doi.org/10.1017/CBO9781139029377.012. Last access on: 13 Mar 2024.

COSTA, A.; CARAMAZZA, A.; SEBASTIAN-GALLES, N. The cognate facilitation effect: implications for models of lexical access. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 26, n. 5, p. 1283-1296, 2000. DOI: https://doi.org/10.1037/0278-7393.26.5.1283. Last access on: 13 Mar 2024.

COSTA, A.; PANNUNZI, M.; DECO, G.; PICKERING, M. J. Do bilinguals automatically activate their native language when they are not using it? **Cognitive Science**, [*S. l.*], v. 41, n. 6, p. 1629-1644, 2017. DOI: https://doi.org/10.1111/cogs.12434. Last access on: 13 Mar 2024.

COTINGUIBA, M. L. P.; COTINGUIBA, G. C.; ANDRETTA, P. I. S. (org.). **Bon Bagay**: Glosè Kreyòl – Ayisyen-Pòtigè – Kreyòl Ayisyen [**Bon Bagay**: Glossário Português-Crioulo Haitiano – Crioulo Haitiano-Português]. Porto Velho, RO: Temática, 2018. ISBN: 978-85-65720-59-5. (*E-book*).

CRISTOFFANINI, P.; KIRSNER, K.; MILECH, D. Bilingual lexical representation: the status of Spanish-English cognates. **The Quarterly Journal of Experimental Psychology Section A**, [*S. l.*], v. 38, n. 3, p. 367-393, 1986. DOI: https://doi.org/10.1080/14640748608401604. Last access on: 13 Mar 2024.

DAHAN, D.; MAGNUSON, J. S. Spoken-word recognition. *In*: TRAXLER, M. J.; GERNSBACHER, M. A. (ed.). **Handbook of Psycholinguistics**. Amsterdam, The Netherlands: Elsevier, 2006. p. 249-283.

DARCY, N. T. A review of the literature on the effects of bilingualism upon the measurement of intelligence. **The Journal of Genetic Psychology**, [*S. l.*], v. 82, n. 1, p. 21-57, 1953. DOI: https://doi.org/10.1080/08856559.1953.10533654. Last access on: 13 Mar 2024.

DAVIS, A. K.; BOWMAN, N.; KAUSHANSKAYA, M. Bilingualism, cognates and reading fluency in children. **Journal of Research in Reading**, [*S. l.*], v. 41, n. S1, p. S12-S29, 2018. DOI: https://doi.org/10.1111/1467-9817.12263. Last access on: 13 Mar 2024.

DAVIS, C. W.; SÁNCHEZ-CASAS, R.; GARCÍA-ALBEA, J. E. **Bilingual lexical representation as revealed using the masked priming procedure**. [*S. l.*: *s. n.*], 1991. Unpublished manuscript.

DE BLESER, R.; DUPONT, P.; POSTLER, J.; BORMANS, G.; SPEELMAN, D.; MORTELMANS, L.; DEBROCK, M. The organisation of the bilingual lexicon: a PET study. **Journal of Neurolinguistics**, [*S. l.*], v. 16, n. 4-5, p. 439-456, 2003. DOI: https://doi.org/10.1016/S0911-6044(03)00022-8. Last access on: 13 Mar 2024.

DE GROOT, A. M. B.; DELMAAR, P.; LUPKER, S. J. The processing of interlexical homographs in translation recognition and lexical decision: support for non-selective access to bilingual memory. **The Quarterly Journal of Experimental Psychology Section A**, [*S. l.*], v. 53, n. 2, p. 397-428, 2000. DOI: https://doi.org/10.1080/713755891. Last access on: 13 Mar 2024.

DE GROOT, A. M. B.; NAS, G. L. Lexical representation of cognates and noncognates in compound bilinguals. **Journal of Memory and Language**, [*S. l.*], v. 30, n. 1, p. 90-123, 1991. DOI: https://doi.org/10.1016/0749-596X(91)90012-9. Last access on: 13 Mar 2024.

DE LEEUW, J.; GILBERT, R. A.; LUCHTERHANDT, B. jsPsych: enabling an open-source collaborative ecosystem of behavioral experiments. **Journal of Open Source Software**, [*S. l.*], v. 8, n. 85, s. p., 2023. DOI: https://doi.org/10.21105/joss.05351. Last access on: 13 Mar 2024.

DE MORAES, I. A.; DE ANDRADE, C. A. A.; MATTOS, B. R. B. A imigração haitiana para o Brasil: causas e desafios. **Conjuntura Austral**, Porto Alegre, v. 4, n. 20, p. 95-114, 2013. DOI: 10.22456/2178-8839.35798. Available at: https://seer.ufrgs.br/index.php/ConjunturaAustral/article/view/35798. Last access on: 13 Mar 2024.

DECLERCK, M.; ÖZBAKAR, E.; KIRK, N. W. Is there proactive inhibitory control during bilingual and bidialectal language production? **PloS One**, [*S. l.*], v. 16, n. 9, e0257355, 2021. DOI: https://doi.org/10.1371/journal.pone.0257355. Last access on: 13 Mar 2024.

DEFIOR, S. Phonological awareness and learning to read: a crosslinguistic perspective. *In*: BRYANT, P.; NUNES, T. (ed.). **Handbook on children's literacy**. London: Academic Press, 2004. p. 631-649.

DEGRAFF, M. Mother-tongue books in Haiti: the power of Kreyòl in learning to read and in reading to learn. **Prospects**, [*S. l.*], v. 46, p. 435-464, 2016.

DEGRAFF, M.; STUMP, G. S. Teaching Linguistics & Language and Public Policy Kreyòl, pedagogy, and technology for opening up quality education in Haiti: changes in teachers' metalinguistic attitudes as first steps in a paradigm shift. **Language**, [*S. l.*], v. 94, n. 2, p. e127-e157, 2018. Available at: http://lingphil.scripts.mit.edu/papers/degraff/DeGraff_and_Stump_2018_Kreyol_pedagogy_and_technology_for_education_in_Haiti.pdf. Last access on: 13 Mar 2024.

DE HOUWER, A.; BORNSTEIN, M. H.; PUTNICK, D. L. A bilingual-monolingual comparison of young children's vocabulary size: evidence from comprehension and production. **Applied Psycholinguistics**, [*S. l.*], v. 35, n. 6, p. 1189-1211, 2014. DOI: https://doi.org/10.1017/S0142716412000744. Last access on: 29 May 2014.

DIDION, L.; TOSTE, J. R.; FILDERMAN, M. J. Teacher professional development and student reading achievement: a meta-analytic review of the effects. **Journal of Research on Educational Effectiveness**, [*S. l.*], v. 13, n. 1, p. 29-66, 2020. DOI: www.doi.org/10.1080/19345747.2019.1670884. Last access on: 13 Mar 2024.

DIJKSTRA, T. Bilingual visual word recognition and lexical access. *In*: KROLL, J. F.; DE GROOT, A. M. B. (ed.). **Handbook of bilingualism**: Psycholinguistic approaches. Oxford: Oxford University Press, 2005. p. 179-201.

DIJKSTRA, T.; GRAINGER, J.; VAN HEUVEN, W. J. B. Recognition of cognates and interlingual homographs: the neglected role of phonology. **Journal of Memory and Language**, [*S. l.*], v. 41, n. 4, p. 496-518, 1999. DOI: https://doi.org/10.1006/jmla.1999.2654. Last access on: 13 Mar 2024.

DIJKSTRA, T.; MIWA, K.; BRUMMELHUIS, B.; SAPPELLI, M.; BAAYEN, H. How cross-language similarity and task demands affect cognate recognition. **Journal of Memory and Language**, [*S. l.*], v. 62, p. 284-301, 2010. DOI: https://doi.org/10.1016/j.jml.2009.12.003. Last access on: 13 Mar 2024.

DIJKSTRA, T.; VAN HEUVEN, W. J. B. The architecture of the bilingual word recognition system: from identification to decision. **Bilingualism**: Language and Cognition, [*S. l.*], v. 5, n. 3, 175-197, 2002. DOI: https://doi.org/10.1017/S1366728902003012. Last access on: 13 Mar 2024.

DIJKSTRA, T.; VAN HEUVEN, W. J. B. The BIA model and bilingual word recognition. *In*: GRAINGER, J.; JACOBS, A. M. (ed.). **Localist connectionist approaches to human cognition**. New Jersey: Lawrence Erlbaum Associates Publishers, 1998. p. 189-225.

DIJKSTRA, T.; VAN JAARSVELD, H.; TEN BRINKE, S. Interlingual homograph recognition: effects of task demands and language intermixing. **Bilingualism**: Language and Cognition, [*S. l.*], v. 1. n. 1, p. 51-66, 1998. DOI: https://doi.org/10.1017/S1366728998000121. Last access on: 13 Mar 2024.

DUFFY, S. A.; MORRIS, R. K.; RAYNER, K. Lexical ambiguity and fixation times in reading. **Journal of Memory and Language**, [*S. l.*], v. 27, n. 4, p. 429-446, 1988. DOI: https://doi.org/10.1016/0749-596X(88)90066-6. Last access on: 13 Mar 2024.

EAMES, T. H. The speed of picture recognition and the speed of word recognition in cases of reading difficulty. **American Journal of Ophthalmology**, [*S. l.*], v. 21, n. 12, p. 1370-1375, 1938. DOI: https://doi.org/10.1016/S0002-9394(38)90409-X. Last access on: 13 Mar 2024.

ECE DEMIR-LIRA, Ö.; APPLEBAUM, L. R.; GOLDIN-MEADOW, S.; LEVINE, S. C. Parents' early book reading to children: relation to children's later language and literacy outcomes controlling for other parent language input. **Developmental Science**, [*S. l.*], v. 22, n. 3, e12764, 2019. DOI: https://doi.org/10.1111/desc.12764. Last access on: 13 Mar 2024.

EGAN, C.; OPPENHEIM, G. M.; SAVILLE, C.; MOLL, K.; JONES, M. W. Bilinguals apply language-specific grain sizes during sentence reading. **Cognition**, [*S. l.*], v. 193, e104018, 2019. DOI: https://doi.org/10.1016/j.cognition.2019.104018. Last access on: 13 Mar 2024.

EHRI, L. C. How children learn to read words. *In*: POLLATSEK, A.; TREIMAN, R. (ed.). **The Oxford handbook of reading**. New York: Oxford University Press, 2015. p. 293-310.

ENGEL DE ABREU, P. M. J.; TOURINHO, C. J.; PUGLISI, M. L.; NIKAEDO, C.; ABREU, N.; MIRANDA, M. C.; BEFI-LOPES, D. M.; BUENO, O. F. A.; MARTIN, R. **A pobreza e a mente**: perspectiva da ciência cognitiva. Walferdange, Luxembourg: The University of Luxembourg, 2015. Available at: https://cienciaparaeducacao.org/wp-content/uploads/2015/10/A-pobreza-e-a-mente_perspectiva-da-ciencia-cognitiva_DEVPOLUX.compressed.pdf. Last access on: 13 Mar 2024.

ERLAM, R.; PHILP, J.; FEICK, D. The adolescent language learner: setting the scene. *In*: ERLAM, R.; PHILP, J.; FEICK, D. **Teaching languages to adolescent learners**: from theory to practice. Cambridge: Cambridge University Press, 2021. p. 1-18. DOI: https://doi.org/10.1017/9781108869812.002. Last access on: 13 Mar 2024.

FEI, X.; ZHAO, S.; LIU, J. Auditory recognition of Chinese-Japanese cognates and homographs by Chinese JFL learners. **Psychologia**, [*S. l.*], v. 64, n. 1, p. 1-22, 2022. DOI: https://doi.org/10.2117/psysoc.2021-A144. Last access on: 13 Mar 2024.

FERREIRA, L. C. Prefácio. *In*: SOUZA, R. D. de; COURA-SOBRINHO, J.; DINIZ, M. B. N. P. (org.). **Português como língua de acolhimento**: práticas e perspectivas. São Paulo: Pá de Palavra, 2021. Available at: https://www.sri.cefetmg.br/wp-content/uploads/sites/84/2022/11/Portugues-como-lingua-de-acolhi-Romulo-Francisco-de-Souza-Jero_compressed.pdf. Last access on: 13 Mar 2024.

FITZ-GERALD, J. D. The bilingual-biracial problem of our border states. **Hispania**, [*S. l.*], v. 4, n. 4, p. 175-186, 1921. DOI: https://doi.org/10.2307/331376. Last access on: 13 Mar 2024.

FISHMAN, J. A. Bilingualism with and without diglossia; diglossia with and without bilingualism. **Journal of Social Issues**, [*S. l.*], v. 23, n. 2, p. 29-38, 1967. DOI: https://doi.org/10.1111/j.1540-4560.1967.tb00573.x. Last access on: 12 Jun 2024.

FLORES, C.; MELO-PFEIFER, S. O conceito "Língua de Herança" na perspetiva da Linguística e da Didática de Línguas: considerações pluridisciplinares em torno do perfil linguístico das crianças lusodescendentes na Alemanha. **Domínios de Lingu@gem**, Uberlândia, v. 8, n. 3, p. 16-45, 2014. DOI: 10.14393/DLesp-v8n3a2014-3. Disponível em: https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/24736. Last access on: 25 May 2024.

FLORIT, E.; BARACHETTI, C.; MAJORANO, M.; LAVELLI, M. Home language activities and expressive vocabulary of toddlers from low-SES monolingual families and bilingual immigrant families. **International Journal of Environmental Research and Public Health**, [*S. l.*], v. 18, n. 1, e296, 2021. DOI: https://doi.org/10.3390/ijerph18010296. Last access on: 13 Mar 2024.

FORCELINI, J.; SUNDERMAN, G. When more is less: the effect of a third language on a second language. **Hispania**, [*S. l.*], v. 103, n. 4, p. 489-500, 2020. DOI: https://www.jstor.org/stable/27026452. Last access on: 13 Mar 2024.

FOWLER, C. A.; MAGNUSON, J. S. Speech perception. *In*: SPIVEY, M. J.; MCRAE, K.; JOANISSE, M. F. (ed.). **The Cambridge handbook of Psycholinguistics**. [*S. l.*]: Cambridge University Press, 2012. p. 3-25. DOI: https://doi.org/10.1017/CBO9781139029377.003. Last access on: 13 Mar 2024.

FREEMAN, M. R.; BLUMENFELD, H. K.; MARIAN, V. Cross-linguistic phonotactic competition and cognitive control in bilinguals. **Journal of Cognitive Psychology**, [*S. l.*], v. 29, n. 7, p. 783-794, 2017. DOI: https://doi.org/10.1080/20445911.2017.1321553. Last access on: 13 Mar 2024.

FREEMAN, M. R.; BLUMENFELD, H. K.; MARIAN, V. Phonotactic constraints are activated across languages in bilinguals. **Frontiers in Psychology**, [*S. l.*], v. 7, e702, 2016. DOI: https://doi.org/10.3389/fpsyg.2016.00702. Last access on: 13 Mar 2024.

FRICKE, M. Modulation of cross-language activation during bilingual auditory word recognition: effects of language experience but not competing background noise. **Frontiers in Psychology**, [*S. l.*], v. 13, e674157, 2022. DOI: https://doi.org/10.3389/fpsyg.2022.674157. Last access on: 13 Mar 2024.

FRITH, U. Beneath the surface of developmental dyslexia. *In*: PATTERSON, K. E.; MARSHALL, J. C.; COLTHEAR, M. (ed.). **Surface dyslexia**: neuropsychological and cognitive analyses of phonological reading. London: Lawrence Erlbaum, 1985. p. 301-330.

FROST, S. J.; LANDI, N.; MENCL, W. E.; SANDAK, R.; FULBRIGHT, R. K.; TEJADA, E. T.; JACOBSEN, L.; GRIGORENKO, E. L.; CONSTABLE, R. T.; PUGH, K. R. Phonological awareness predicts activation patterns for print and speech. **Annals of Dyslexia**, [*S. l.*], v. 59, p. 78-97, 2009. DOI: https://doi.org/10.1007/s11881-009-0024-y. Last access on: 13 Mar 2024.

GARCÍA-ALBEA, J. E.; SÁNCHEZ-CASAS, R.; IGOA, J. M. The contribution of word form and meaning to language processing in Spanish: some evidence from monolingual and bilingual studies. *In*: HILLERT, D. (ed.). **Sentence processing**: a cross-linguistic perspective. New York: Academic Press, 1998. p. 183-209.

GARCÍA-ALBEA, J. E.; SÁNCHEZ-CASAS, R.; VALERO, T. Form and meaning contribution to word recognition in Catalan-Spanish bilinguals. Paper presented at the meeting of the **Ninth Conference of the European Society for Cognitive Psychology**, University of Würzburg, Germany, 1996.

GASTMANN, F.; POARCH, G. J. Cross-language activation during word recognition in child second-language learners and the role of executive function. **Journal of Experimental Child Psychology**, [*S. l.*], v. 221, e105443, 2022. DOI: https://doi.org/10.1016/j.jecp.2022.105443. Last access on: 13 Mar 2024.

GAVINO, M. F.; GOLDRICK, M. Bilingual language processing: interactions between lexical retrieval and vowel production. **The Journal of the Acoustical Society of America**, [*S. l.*], v. 148, e2657, 2020. DOI: https://doi.org/10.1121/1.5147401. Last access on: 13 Mar 2024.

GERARD, L. D.; SCARBOROUGH, D. L. Language-specific lexical access of homographs by bilinguals. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 15, n. 2, p. 305-315, 1989. DOI: https://doi.org/10.1037/0278-7393.15.2.305. Last access on: 13 Mar 2024.

GERTH, S.; FESTMAN, J. Reading development, word length and frequency effects: an eye-tracking study with slow and fast readers. **Frontiers in Communication**, Section Psychology of Language, [*S. l.*], v. 6, p. 1-12, 2021. DOI: https://doi.org/10.3389/fcomm.2021.743113. Last access on: 13 Mar 2024.

GLAUDE, H.; ZRIBI-HERTZ, A. Verb cognates in Haitian Creole. **The Linguistic Review**, [*S. l.*], v. 31, n. 2, p. 241-294, 2014. DOI: https://doi.org/10.1515/tlr-2014-0003. Last access on: 13 Mar 2024.

GOMES, M. A. Os impactos subjetivos dos fluxos migratórios: os haitianos em Florianópolis (SC). **Psicologia & Sociedade**, Belo Horizonte, v. 29, e162484, 2017. DOI: https://doi.org/10.1590/1807-0310/2017v29162484. Last access on: 13 Mar 2024.

GOUGH, P. B.; TUNMER, W. E. Decoding, reading, and reading disability. **Remedial and Special Education**, [*S. l.*], v. 7, n. 1, p. 6-10, 1986. DOI: https://doi.org/10.1177/074193258600700104. Last access on: 13 Mar 2024.

GRAINGER, J.; DIJKSTRA, T. On the representation and use of language information in bilinguals. *In*: HARRIS, R. J. (ed.). **Cognitive processing in bilinguals**. Manhattan: North-Holland, 1992. p. 207-220. DOI: https://doi.org/10.1016/S0166-4115(08)61496-X. Last access on: 13 Mar 2024.

GRAINGER, J.; O'REGAN, J. K.; JACOBS, A. M.; SEGUI, J. On the role of competing word units in visual word recognition: the neighborhood frequency effect. **Perception & Psychophysics**, [*S. l.*], v. 45, p. 189-195, 1989. DOI: https://doi.org/10.3758/BF03210696. Last access on: 13 Mar 2024.

GREEN, D. W. Mental control of the bilingual lexico-semantic system. **Bilingualism**: Language and Cognition, [*S. l.*], v. 1, n. 2, p. 67-81, 1998. DOI: 10.1017/S1366728998000133. Last access on: 13 Mar 2024.

GROJEAN, F. Who is Bilingual? *In*: GROSJEAN, F. **The mysteries of bilingualism**: unresolved issues. Chichester: Wiley-Blackwell, 2022. p. 7-17. Available at: https://www.francoisgrosjean.ch/Mysteries_of_Biligualism_ch1.pdf. Last access on: 13 Mar 2024.

GROSJEAN, F. Bilingualism: a short introduction. *In*: GROSJEAN, F.; LI, P. **The Psycholinguistics of bilingualism**. Hoboken: Blackwell Publishing, 2013. p. 5-26.

GROSJEAN, F. Exploring the recognition of guest words in bilingual speech. **Language and Cognitive Processes**, [*S. l.*], v. 3, n. 3, p. 233-274, 1988. DOI: https://doi.org/10.1080/01690968808402089. Last access on: 13 Mar 2024.

GROSJEAN, F. **Studying bilinguals**. New York: Oxford University Press, 2008.

GROSJEAN, F. The extent of bilingualism. *In*: GROSJEAN, F. **Life as a bilingual**: knowing and using two or more languages. Cambridge: Cambridge Press, 2021. p. 27-39. DOI: https://doi.org/10.1017/9781108975490.003. Last access on: 13 Mar 2024.

GROSJEAN, F.; LI, P. **The Psycholinguistics of bilingualism**. Hoboken: John Wiley & Sons, 2013.

GROSS, J.; LIGGENS, U. **Tests of normality**. Package "nortest". [*S. l.*: *s. n.*]: 2015. Available at: https://cran.r-project.org/web/packages/nortest/nortest.pdf. Last access on: 13 Mar 2024.

GUAGLIANO, C. Florianópolis se torna 2º município no Brasil a ter Política Municipal para a População Migrante. **MigraMundo**, [*S. l.*], 6 Ago 2020. Available at: https://migramundo.com/florianopolis-se-torna-2o-municipio-no-brasil-a-ter-politica-municipal-para-a-populacao-migrante/. Last access on: 13 Mar 2024.

GUASCH, M. **Forma y significado en el procesamiento léxico de bilingües del castellano y del catalán**. 2001. Thesis (Master's Degree) – Universitat Rovira i Virgili, Tarragona, Spain, 2001. Unpublished.

GUASCH, M.; BOADA, R.; FERRÉ, P.; SÁNCHEZ-CASAS, R. NIM: a web-based Swiss army knife to select stimuli for psycholinguistic studies. **Behavior Research Methods**, [*S. l.*], v. 45, p. 765-771, 2013. DOI: https://doi.org/10.3758/s13428-012-0296-8. Last access on: 13 Mar 2024.

GULLIFER, J. W.; KROLL, J. F.; DUSSIAS, P. E. When language switching has no apparent cost: lexical access in sentence context. **Frontiers in Psychology**, [*S. l.*], v. 4, e278, 2013. DOI: https://doi.org/10.3389/fpsyg.2013.00278. Last access on: 13 Mar 2024.

HAIGH, C. A.; JARED, D. The activation of phonological representations by bilinguals while reading silently: evidence from interlingual homophones. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 33, n. 4, p. 623-644, 2007. DOI: https://doi.org/10.1037/0278-7393.33.4.623. Last access on: 13 Mar 2024.

HAMEAU, S.; BIEDERMANN, B.; NICKELS, L. Lexical activation in late bilinguals: effects of phonological neighbourhood on spoken word production. **Language, Cognition and Neuroscience**, [*S. l.*], v. 36, n. 4, p. 517-534, 2021. DOI: https://doi.org/10.1080/23273798.2020.1863438. Last access on: 13 Mar 2024.

HARM, M. W.; SEIDENBERG, M. S. Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. **Psychological Review**, [*S. l.*], v. 111, n. 3, p. 662-720, 2004. DOI: https://doi.org/10.1037/0033-295X.111.3.662. Last access on: 13 Mar 2024.

HASSAN, I.; AZMI, M. N.L.; MUHAMAD, S. N.; ABDULLAH, A. T. H. Reading habits and their correlation with reading achievement among ESL learners in selected Malaysian secondary schools. **Arab World English Journal**, [*S. l.*], v. 12, n. 3, p. 385-399, 2021. DOI: https://dx.doi.org/10.24093/awej/vol12no3.27. Last access on: 13 Mar 2024.

HATCHER, P. J.; HULME, C.; ELLIS, A. W. Ameliorating early reading failure by integrating the teaching of reading and phonological skills: the phonological linkage hypothesis. **Child Development**, [*S. l.*], v. 65, n. 1, p. 41-57, 1994. DOI: https://doi.org/10.2307/1131364. Last access on: 13 Mar 2024.

HAUGEN, E. Problems of bilingualism. **Lingua**, [*S. l.*], v. 2, p. 271-290, 1949. DOI: https://doi.org/10.1016/0024-3841(49)90028-5. Last access on: 13 Mar 2024.

HEBBLETHWAITE, B. Adverb code-switching among Miami's Haitian Creole-English second generation. **Bilingualism**: Language and Cognition, [*S. l.*], v. 13, n. 4, p. 409-428, 2010. DOI: 10.1017/S1366728909990563. Last access on: 13 Mar 2024.

HELMS-PARK, R.; PERHAN, Z. The role of explicit instruction in cross-script cognate recognition: the case of Ukrainian-speaking EAP learners. **Journal of English for Academic Purposes**, [*S. l.*], v. 21, p. 17-33, 2016. DOI: https://doi.org/10.1016/j.jeap.2015.08.005. Last access on: 13 Mar 2024.

HENRICH, J.; HEINE, S. J.; NORENZAYAN, A. The weirdest people in the world? **The Behavioral and Brain Sciences**, [*S. l.*], v. 33, n. 2-3, p. 61-135, 2010. DOI: https://doi.org/10.1017/S0140525X0999152X. Last access on: 13 Mar 2024.

HIPFNER-BOUCHER, K.; MILBURN, T.; WEITZMAN, E.; GREENBERG, J.; PELLETIER, J.; GIROLAMETTO, L. Relationships between preschoolers' oral language and phonological awareness. **First Language**, [*S. l.*], v. 34, n. 2, p. 178-197, 2014. DOI: https://doi.org/10.1177/0142723714525945. Last access on: 13 Mar 2024.

HOGAN, T.; CATTS, H.; LITTLE T. The relationship between phonological awareness and reading. **Language, Speech, and Hearing Services in Schools**, [*S. l.*], v. 36, n. 4, p. 285-293, 2005. DOI: https://doi.org/10.1044/0161-1461(2005/029). Last access on: 13 Mar 2024.

HOOVER, W. A.; GOUGH, P. B. The simple view of reading. **Reading and Writing**: An Interdisciplinary Journal, [*S. l.*], v. 2, n. 2, p. 127-160, 1990. DOI: https://doi.org/10.1007/BF00401799. Last access on: 13 Mar 2024.

HOSHINO, N.; KROLL, J. F. Cognate effects in picture naming: does cross-language activation survive a change of script? **Cognition**, [*S. l.*], v. 106, n. 1, p. 501-511, 2008. DOI: https://doi.org/10.1016/j.cognition.2007.02.001. Last access on: 13 Mar 2024.

HSIEH, M.-C.; JEONG, H.; KAWATA, K. H. dos S.; SASAKI, Y.; LEE, H.-C.; YOKOYAMA, S.; SUGIURA, M.; KAWASHIMA, R. Neural correlates of bilingual language control during interlingual homograph processing in a logogram writing system. **Brain and Language**, [*S. l.*], v. 174, p. 72-85, 2017. DOI: https://doi.org/10.1016/j.bandl.2017.06.006. Last access on: 13 Mar 2024.

HSIEH, M.-C.; JEONG, H.; SUGIURA, M.; KAWASHIMA, R. Neural evidence of language membership control in bilingual word recognition: an fMRI study of cognate processing in Chinese-Japanese Bilinguals. **Frontiers in Psychology**, [*S. l.*], v. 12, p. 1-14, 2021. DOI: https://doi.org/10.3389/fpsyg.2021.643211.

HÜBNER, L. C. *et al.* Bateria de Avaliação da Linguagem no Envelhecimento – BALE. *In*: ZIMMERMANN, N.; DELEAERE, F.; FONSECA, R. P. (org.). **Tarefas de avaliação neuropsicológica para idosos**. Rio de Janeiro: Memnon, 2019.

IBÁÑEZ-ALFONSO, J. A.; HERNÁNDEZ-CABRERA, J. A.; DUÑABEITIA, J. A.; ESTÉVEZ, A.; MACIZO, P.; BAJO, M. T.; FUENTES, L. J.; SALDAÑA, D. Socioeconomic status, culture, and reading comprehension in immigrant students. **Frontiers in Psychology**, [*S. l.*], v. 12, 752273, 2021. DOI: https://doi.org/10.3389/fpsyg.2021.752273. Last access on: 13 Mar 2024.

IMDH – INSTITUTO MIGRAÇÕES E DIREITOS HUMANOS. Haitianos no Brasil: dados estatísticos, informações e uma recomendação. **Instituto Migrações e Direitos Humanos**, [*S. l.*], 2016. Available at: https://www.migrante.org.br/wp-content/uploads/2016/12/Haitianos_dados-PF_CNIg-Recomendao-de-Registro-1.pdf. Last access on: 2 Jul 2024.

INIESTA, A.; ROSSI, E.; BAJO, M. T.; PAOLIERI, D. The influence of cross-linguistic similarity and language background on writing to dictation. **Frontiers in Psychology**, [*S. l.*], v. 12, e679956, 2021. DOI: https://doi.org/10.3389/fpsyg.2021.679956. Last access on: 13 Mar 2024.

IOM – INTERNATIONAL ORGANIZATION FOR MIGRATION. **Brazil visa application centre**. [*S. l.*]: IOM, 2023. Available at: https://www.iom.int/sites/g/files/tmzbdl486/files/documents/2023-07/brazil-humanitarian-visa-support-centre.pdf. Last access on: 13 Mar 2024.

IPHAN – INSTITUTO DO PATRIMÓNIO HISTÓRICO E ARTÍSTICO NACIONAL. Diversidade Linguística – No Brasil, são faladas mais de 250 línguas. **Inventário Nacional da Diversidade Linguística (INDL)**, [*S. l.*], 2014. Available at: http://portal.iphan.gov.br/indl. Last access on: 13 Mar 2024.

IPOL – INSTITUTO DE INVESTIGAÇÃO E DESENVOLVIMENTO EM POLÍTICA LINGUÍSTICA. Sobre o IPOL. **IPOL**, Florianópolis, [20-?]. Available at: http://ipol.org.br/sobre-o-ipol/. Last access on: 13 Mar 2024.

JACKSON, C. N.; MASSARO, A.; HOPP, H. The impact of L1 structural frequency and cognate status on the timing of L2 production. **Journal of Cognitive Psychology**, [*S. l.*], v. 29, n. 5, p. 535-550, 2017. DOI: https://doi.org/10.1080/20445911.2017.1299156. Last access on: 13 Mar 2024.

JACOBS, A.; FRICKE, M.; KROLL, J. F. Cross-language activation begins during speech planning and extends into second language speech. **Language Learning**, [*S. l.*], v. 66, n. 2, p. 324-353, 2016. DOI: https://doi.org/10.1111/lang.12148. Last access on: 13 Mar 2024.

JARED, D. Literacy and literacy development in bilinguals. *In*: POLLATSEK, A.; TREIMAN, R. **The Oxford handbook of reading**. New York: Oxford University Press, 2015. p. 165-182.

JARED, D.; CORMIER, P.; LEVY, B. A.; WADE-WOOLLEY, L. Cross-language activation of phonology in young bilingual readers. **Reading and Writing**: An Interdisciplinary Journal, [*S. l.*], v. 25, n. 6, p. 1327-1343, 2012. DOI: https://doi.org/10.1007/s11145-011-9320-0. Last access on: 13 Mar 2024.

JOSS, L. M.; VIRTUE, S. Hemispheric processing in bilinguals: the role of shared meanings across languages and sentential constraint. **Cognitive Neuroscience**, [*S. l.*], v. 1, n. 1, p. 26-32, 2010. DOI: https://doi.org/10.1080/17588920903420167. Last access on: 13 Mar 2024.

JUSTI, F. R. dos R.; JUSTI, C. N. G.; ROAZZI, A. Efeitos da similaridade ortográfica das pseudopalavras no acesso lexical. **Arquivos Brasileiros de Psicologia**, [*S. l.*], v. 66, n. 3, p. 133-147, 2014. Available at: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1809-52672014000300011&lng=pt&nrm=iso. Last access on: 13 Mar 2024.

KAZEMI, M.; MEHRI, A.; JALAEI, S.; GOLZADEH, F. A. The study of cognate and non-cognate nouns and verbs naming in Mazandarani-Persian bilinguals. **Journal of Modern Rehabilitation**, [*S. l.*], v. 17, n. 1, 2022. DOI: https://doi.org/10.18502/jmr.v17i1.11305. Last access on: 13 Mar 2024.

KENNISON, S. M.; FERNANDEZ, E. C.; BOWERS, J. M. The roles of semantic and phonological information in word production: evidence from Spanish-English bilinguals. **Journal of Psycholinguistic Research**, [*S. l.*], v. 43, n. 2, p. 105-124, 2014. DOI: https://doi.org/10.1007/s10936-013-9243-9. Last access on: 13 Mar 2024.

KESSLER, B.; TREIMAN, R. Writing systems: their properties and implications for reading. *In*: POLLATSEK, A.; TREIMAN, R. **The Oxford handbook of reading**. New York: Oxford University Press, 2015. p. 10-26.

KIM, J.; DAVIS, C. Task effects in masked cross-script translation and phonological priming. **Journal of Memory and Language**, [*S. l.*], v. 49, n. 4, p. 484-499, 2003. DOI: https://doi.org/10.1016/S0749-596X(03)00093-7. Last access on: 13 Mar 2024.

KIRK, N. W.; KEMPE, V.; SCOTT-BROWN, K. C.; PHILIPP, A.; DECLERCK, M. Can monolinguals be like bilinguals? Evidence from dialect switching. **Cognition**, [*S. l.*], v. 170, p. 164-178, 2018. DOI: https://doi.org/10.1016/j.cognition.2017.10.001. Last access on: 13 Mar 2024.

KLEINEBING, F. Secretaria de Estado da Educação promove capacitação sobre o Programa de Acolhimento ao Migrante e Refugiado (PARE). **sc.gov.br**, Florianópolis, 11 Ago. 2023. Available at: https://estado.sc.gov.br/noticias/secretaria-de-estado-da-educacao-promove-capacitacao-sobre-o-programa-de-acolhimento-ao-migrante-e-refugiado-pare/. Last access on: 13 Mar 2024.

KOOTSTRA, G. J.; DIJKSTRA, T.; VAN HELL, J. G. Interactive alignment and lexical triggering of code-switching in bilingual dialogue. **Frontiers in Psychology**, [*S. l.*], v. 11, e1747, 2020. DOI: https://doi.org/10.3389/fpsyg.2020.01747. Last access on: 13 Mar 2024.

KORMOS, J.; KIDDLE, T. The of socio-economic factors in motivation to learn English as a foreign language: the case of Chile. **System**, [*S. l.*], v. 41, p. 399-412, 2013. DOI: http://dx.doi.org/10.1016/j.system.2013.03.006. Last access on: 13 Mar 2024.

KRIEGEL, S. Haitian Creole: structure, variation, status, origin by Albert Valdman (review). **Anthropological Linguistics**, [*S. l.*], v. 58, n. 3, p. 329-332, 2016. DOI: https://doi.org/10.1353/anl.2016.0034. Last access on: 13 Mar 2024.

KROGH, S. M. Danish-English bilinguals' cognate processing in L1 and L2 visual lexical decision tasks. **Languages**, [*S. l.*], v. 7, n. 3, e228, 2022. DOI: https://doi.org/10.3390/languages7030228. Last access on: 13 Mar 2024.

KROLL, J. F.; DUSSIAS, P. E.; BICE, K.; PERROTTI, L. Bilingualism, mind, and brain. **Annual Review of Linguistics**, [*S. l.*], v. 1, p. 377-394, 2015. DOI: https://doi.org/10.1146/annurev-linguist-030514-124937. Last access on: 13 Mar 2024.

KROLL, J. F.; MA, F. Y. The bilingual lexicon. *In*: FERNÁNDEZ, E. M.; CAIRNS, H. S. (ed.). **The handbook of Psycholinguistics**. Hoboken: John Wiley & Sons, 2018. p. 294-319. DOI: https://doi.org/10.1002/9781118829516.ch13. Last access on: 13 Mar 2024.

KROLL, J. F.; STEWART, E. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. **Journal of Memory and Language**, [*S. l.*], v. 33, p. 149-174, 1994. DOI: https://doi.org/10.1006/jmla.1994.1008. Last access on: 13 Mar 2024.

KROLL, J. F.; VAN HELL, J. G.; TOKOWICZ, N.; GREEN, D. W. The revised hierarchical model: a critical review and assessment. **Bilingualism**, [*S. l.*], v. 13, n. 3, p. 373-381, 2010. DOI: https://doi.org/10.1017/S136672891000009X. Last access on: 13 Mar 2024.

KUHL, P. K.; CONBOY, B. T.; COFFEY-CORINA, S.; PADDEN, D.; RIVERA-GAXIOLA, M.; NELSON, T. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). **Philosophical Transactions of the Royal Society B**: Biological Sciences, [*S. l.*], v. 363, n. 1493, p. 979-1000, p. 2007. DOI: https://doi.org/10.1098/rstb.2007.2154. Last access on: 29 may 2024.

KUZNETSOVA, A.; BROCKHOFF, P. B.; CHRISTENSEN, R. H. B. lmerTest package: tests in linear mixed effects models. **Journal of Statistical Software**, [*S. l.*], v. 82, n. 13, p. 1-26, 2017.

LALOR, E.; KIRSNER, K. The representation of "false cognates" in the bilingual lexicon. **Psychonomic Bulletin & Review**, [*S. l.*], v. 8, n. 3, p. 552-559, 2001. DOI: https://doi.org/10.3758/bf03196191. Last access on: 13 Mar 2024.

LAURO, J.; SCHWARTZ, A. I. Bilingual non-selective lexical access in sentence contexts: a meta-analytic review. **Journal of Memory and Language**, [*S. l.*], v. 92, p. 217-233, 2017. DOI: https://doi.org/10.1016/j.jml.2016.06.010. Last access on: 13 Mar 2024.

LEBEL, C.; WALKER, L.; LEEMANS, A.; PHILLIPS, L.; BEAULIEU, C. Microstructural maturation of the human brain from childhood to adulthood. **NeuroImage**, [*S. l.*], v. 40, n. 3, p. 1044-1055, 2008. DOI: https://doi.org/10.1016/j.neuroimage.2007.12.053. Last access on: 13 Mar 2024.

LEFEBVRE, C. **Creole genesis and the acquisition of grammar**: the case of Haitian Creole. New York: Cambridge University Press, 1998. Cambridge Studies in Linguistics.

LEFEBVRE, C. On data. **Journal of Pidgin and Creole Languages**, [*S. l.*], v. 15, n. 2, p. 313-337, 1997. DOI: https://doi.org/10.1075/jpcl.15.2.04lef. Last access on: 13 Mar 2024.

LÉGER, F.; ARMBRISTER, A. P. Factors affecting the teaching and learning of Haitian Creole in the Bahamas. **The College of The Bahamas Research Journal**, [*S. l.*], v. 15, p. 22-35, 2009.

LEIPZIG UNIVERSITY. Wikipedia 2021, Haitian. **Corpora Collection Leipzig**, [*S. l.*], [2024]. Available at: https://corpora.uni-leipzig.de/en?corpusId=hat_wikipedia_2021. Last access on: 13 Mar 2024.

LEIVADA, E.; RODRÍGUEZ-ORDÓÑEZ, I.; PARAFITA COUTO, M. C.; PERPIÑÁN, S. Bilingualism with minority languages: why searching for unicorn language users does not move us forward. **Applied Psycholinguistics**, [*S. l*], v. 44, n. 3, p. 384-399, 2023. DOI: https://doi.org/10.1017/S0142716423000036. Last access on: 12 jun. 2024.

LEMHÖFER, K.; DIJKSTRA, T. Recognizing cognates and interlingual homographs: effects of code similarity in language-specific and generalized lexical decision. **Memory & Cognition**, [*S. l.*], v. 32, n. 4, p. 533-550, 2004. DOI: https://doi.org/10.3758/BF03195845. Last access on: 13 Mar 2024.

LEMHÖFER, K.; DIJKSTRA, T.; SCHRIEFERS, H.; BAAYEN, R. H.; GRAINGER, J.; ZWITSERLOOD, P. Native language influences on word recognition in a second language: a megastudy. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 34, n. 1, p. 12-31, 2008. DOI: https://doi.org/10.1037/0278-7393.34.1.12. Last access on: 13 Mar 2024.

LEMHÖFER, K.; HUESTEGGE, L.; MULDER, K. Another cup of TEE? The processing of second language near-cognates in first language reading. **Language, Cognition and Neuroscience**, [*S. l.*], v. 33, n. 8, p. 968-991, 2018. DOI: https://doi.org/10.1080/23273798.2018.1433863. Last access on: 13 Mar 2024.

LEMHÖFER, K.; SPALEK, K.; SCHRIEFERS, H. Cross-language effects of grammatical gender in bilingual word recognition and production. **Journal of Memory and Language**, [*S. l.*], v. 59, n. 3, p. 312-330, 2008. DOI: https://doi.org/10.1016/j.jml.2008.06.005. Last access on: 13 Mar 2024.

LEREA, L.; KOHUT, S. A comparative study of monolinguals and bilinguals in a verbal task performance. **Journal of Clinical Psychology**, [*S. l.*], v. 17, p. 49-52, 1961. DOI: https://doi.org/10.1002/1097-4679(196101)17:1<49::aid-jclp2270170117>3.0.co;2-n. Last access on: 13 Mar 2024.

LERVÅG, A.; AUKRUST, V. G. Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. **Journal of Child Psychology and Psychiatry, and Allied Sisciplines**, [*S. l.*], v. 51, n. 5, p. 612-620, 2010. DOI: https://doi.org/10.1111/j.1469-7610.2009.02185.x. Last access on: 13 Mar 2024.

LEWELLEN, M. J.; GOLDINGER, S. D.; PISONI, D. B.; GREENE, B. G. Lexical familiarity and processing efficiency: individual differences in naming, lexical decision, and semantic categorization. **Journal of Experimental Psychology**, [*S. l.*], v. 122, n. 3, p. 316-330, 1993. DOI: https://doi.org/10.1037//0096-3445.122.3.316. Last access on: 13 Mar 2024.

LI, C.; GOLLAN, T. H. Cognates facilitate switches and then confusion: contrasting effects of cascade versus feedback on language selection. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 44, n. 6, p. 974-991, 2018a. DOI: https://doi.org/10.1037/xlm0000497. Last access on: 13 Mar 2024.

LI, C.; GOLLAN, T. H. Cognates interfere with language selection but enhance monitoring in connected speech. **Memory & Cognition**, [*S. l.*], v. 46, n. 6, p. 923-939, 2018b. DOI: https://doi.org/10.3758/s13421-018-0812-x. Last access on: 13 Mar 2024.

LI, C.; GOLLAN, T. H. What cognates reveal about default language selection in bilingual sentence production. **Journal of Memory and Language**, [*S. l.*], v. 118, e104214, 2021. DOI: https://doi.org/10.1016/j.jml.2020.104214. Last access on: 13 Mar 2024.

LI, P.; ZHANG, F.; YU, A.; ZHAO, X. Language History Questionnaire (LHQ3): an enhanced tool for assessing multilingual experience. **Bilingualism**: Language and Cognition, [*S. l.*], v. 23, n. 5, p. 1-7, 2019. DOI: https://doi.org/10.1017/S1366728918001153. Last access on: 13 Mar 2024.

LI, P.; ZHAO, X. Self-organizing map models of language acquisition. **Frontiers in Psychology**, [*S. l.*], v. 4, e828, 2013. DOI: https://doi.org/10.3389/fpsyg.2013.00828. Last access on: 13 Mar 2024.

LI, P.; ZHAO, X.; MACWHINNEY, B. Dynamic self-organization and early lexical development in children. **Cognitive Science**, [*S. l.*], v. 31, n. 4, p. 581-612, 2007. DOI: https://doi.org/10.1080/15326900701399905. Last access on: 13 Mar 2024.

LIBBEN, M. R.; TITONE, D. A. Bilingual lexical access in context: evidence from eye movements during reading. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 35, n. 2, p. 381-390, 2009. DOI: https://doi.org/10.1037/a0014875. Last access on: 13 Mar 2024.

LIBERMAN, A. M.; COOPER, F. S.; STUDDERT-KENNEDY, M.; HARRIS, K. S.; SHANKWEILER, D. P. On the efficiency of speech sounds. **Language Typology and Universals**, [*S. l.*], v. 21, n. 1-6, p. 21-32, 1968. DOI: https://doi.org/10.1524/stuf.1968.21.16.21. Last access on: 13 Mar 2024.

LIJEWSKA, A. The influence of semantic bias on triple non-identical cognates during reading: evidence from trilinguals' eye movements. **Second Language Research**, [*S. l.*], v. 39, n. 4, p. 1235-1263, 2023. DOI: https://doi.org/10.1177/02676583221128525. Last access on: 13 Mar 2024.

LIJEWSKA, A.; CHMIEL, A. Cognate facilitation in sentence context – translation production by interpreting trainees and non-interpreting trilinguals. **International Journal of Multilingualism**, [*S. l.*], v. 12, n. 3, p. 358-375, 2015. DOI: https://doi.org/10.1080/14790718.2014.959961. Last access on: 13 Mar 2024.

LIM, J. H.; CHRISTIANSON, K. Cross-script L1-L2 and L2-L1 masked translation priming and phonological priming: evidence from unbalanced Korean-English bilinguals. **International Journal of Bilingualism**, [*S. l.*], v. 27, n. 5, p. 862-881, 2023. DOI: https://doi.org/10.1177/13670069221142405. Last access on: 13 Mar 2024.

LIMA, B.; PERNA, Y. Após bloqueio no MEC, reitores de universidades dizem que terão de "cortar no osso": limpeza, restaurante, luz, água e bolsas estudantis. **G1**, Economia, [*S. l.*], 6 Oct 2022. Available at: https://g1.globo.com/economia/de-olho-no-orcamento/noticia/2022/10/06/para-entidade-corte-do-governo-bolsonaro-na-educacao-afeta-agua-e-luz-e-inviabiliza-universidades.ghtml. Last access on: 13 Mar 2024.

LIMA, J. B. B. (ed.). **Brazilian cooperation for international development**: 2011-2013. Brasília: IPEA; ABC, 2017. Available at: http://www.abc.gov.br/imprensa/mostrarconteudo/900. Last access on: 13 Mar 2024.

LIMA, R. A. S. C. **Tradução, adaptação e validação do test de vocabulario en Imágenes Peabody (TVIP)**. 2007. Tese (Doutorado) – Programa de Pós-Graduação em Psicologia Cognitiva, Universidade Federal de Pernambuco, Recife, 2007.

LIMBERGER, B. K. **Processamento da leitura multilíngue e suas bases neurais**: um estudo sobre o hunsriqueano. 2018. 270 p. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Letras, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2018. Available at: https://tede2.pucrs.br/tede2/handle/tede/7952. Last access on: 30 May 2024.

LOMICKA, L.; LORD, G. Social networking and language learning. *In*: FARR, F.; MURRAY, L. (ed.). **The Routledge handbook of language learning and technology**. New York: Routledge, 2016. p. 255-268.

LOPEZ, A. P. de A. Portuguese as a welcoming language for forcibly displaced immigrants in Brazil: some principles for teaching in light of interculturality. **Revista Brasileira de Linguística Aplicada**, Belo Horizonte, v. 18, n. 2, p. 389-416, 2018. DOI: https://doi.org/10.1590/1984-6398201812048. Last access on: 13 Mar 2024.

LUKATELA, G.; TURVEY, M. T. An evaluation of the two-cycles model of phonology assembly. **Journal of Memory and Language**, [*S. l.*], v. 42, n. 2, p. 183-207, 2000. DOI: https://doi.org/10.1006/jmla.1999.2672. Last access on: 13 Mar 2024.

MACKEY, W. F. Bilingualism in the Americas: a bibliography and research guide, by Einar Haugen. Publication of the American Dialect Society, n. 26, 1956, 159 p. Obtainable from the University of Alabama Press, Drawer 7877. **Canadian Journal of Linguistics/Revue canadienne de linguistique**, [*S. l.*], v. 4, n. 2, p. 94-98, 1958. DOI: https://doi.org/10.1017/S0008413100025470. Last access on: 13 Mar 2024.

MAGALHÃES, L. F. A.; BAENINGER, R. Imigração haitiana no estado de Santa Catarina: fases do fluxo e contradições da inserção laboral. **Geografias**, [*S. l.*], p. 100-111, 2017. DOI: 10.35699/2237-549X.13450. Available at: https://periodicos.ufmg.br/index.php/geografias/article/view/13450. Last access on: 12 Oct 2023.

MANN, V. A.; FOY, J. G. Phonological awareness, speech development, and letter knowledge in preschool children. **Annals of Dyslexia**, [*S. l.*], v. 53, p. 149-173, 2003. DOI: https://doi.org/10.1007/s11881-003-0008-2. Last access: 27 May 2024.

MARCOTTE, K.; ANSALDO, A. I. Age-related behavioural and neurofunctional patterns of second language word learning: different ways of being successful. **Brain and Language**, [*S. l.*], v. 135, p. 9-19, 2014. DOI: https://doi.org/10.1016/j.bandl.2014.04.004. Last access on: 13 Mar 2024.

MARIAN, V.; SPIVEY, M. Bilingual and monolingual processing of competing lexical items. **Applied Psycholinguistics**, [*S. l.*], v. 24, n. 2, p. 173-193, 2003. DOI: https://doi.org/10.1017/S0142716403000092. Last access on: 13 Mar 2024.

MARINIS, T. Using on-line processing methods in language acquisition research. *In*: UNSWORTH, S.; BLOM, E. (ed.). **Experimental methods in language acquisition research**. [*S. l.*]: John Benjamins, 2010. p. 139-162.

MARSH, G.; DESBERG, P. Development of strategies in the acquisition of symbolic skills. *In*: ROGERS, D. R.; SLOBODA, J. A. (ed.). **The acquisition of symbolic skills**. Boston: Plenum, 1983. p. 149-154.

MARSH, G.; DESBERG, P.; COOPER, J. Developmental changes in strategies of reading. **Journal of Reading Behaviour**, [*S. l.*], v. 9, p. 391-394, 1977.

MARSH, G.; FRIEDMAN, M. P.; DESBERG, P.; WELCH, V. Development of strategies in learning to spell. *In*: FRITH, U. (ed.). **Cognitive processes in spelling**. London: Academic Press, 1980. p. 339–353.

MARSH, G.; FRIEDMAN, M. P.; WELCH, V.; DESBERG, P. A cognitive-developmental theory of reading acquisition. *In*: WALLER, T. G.; MACKINNON, G. E. (ed.). **Reading research advances in theory and practice**. New York: Academic Press, 1981. v. 3.

MARTIN-CHANG, S.; OUELLETTE, G.; MADDEN, M. Does poor spelling equate to slow reading? The relationship between reading, spelling, and orthographic quality. **Reading and Writing**: An Interdisciplinary Journal, [*S. l.*], v. 27, n. 8, p. 1485-1505, 2014. DOI: https://doi.org/10.1007/s11145-014-9502-7. Last access on: 13 Mar 2024.

MARTÍNEZ-GARCÍA, M. T. Using eye-movements to track bilingual activation. **Languages**, [*S. l.*], v. 4, n. 3, p. 59, 2019. DOI: https://doi.org/10.3390/languages4030059. Last access on: 13 Mar 2024.

MARTÍNEZ-GARCIA, M. T.; TREMBLAY, A. Effects of language bias and proficiency on cross-language activation: evidence from eye-tracking. **The Journal of the Acoustical Society of America**, [*S. l.*], v. 140, e3341, 2016. DOI: https://doi.org/10.1121/1.4970666. Last access on: 13 Mar 2024.

MAYOR, J.; PLUNKETT, K. A statistical estimate of infant and toddler vocabulary size from CDI analysis. **Developmental Science**, [*S. l.*], v. 14, n. 4, p. 769-785, 2011. DOI: https://doi.org/10.2307/25057972?ref=no-x-route:58

a0ca288783055a71a92e8d3b06ac76. Last access on: 29 May 2014.

MCCLELLAND, J. L.; ELMAN, J. L. The TRACE model of speech perception. **Cognitive Psychology**, [*S. l.*], v. 18, n. 1, p. 1-86, 1986. DOI: https://doi.org/10.1016/0010-0285(86)90015-0. Last access on: 13 Mar 2024.

MCCLELLAND, J. L.; RUMELHART, D. E. An interactive activation model of context effects in letter perception. Part I: An account of basic findings. **Psychological Review**, [*S. l.*], v. 88, p. 375-405, 1981.

MIAO, R. Second language acquisition: an introduction. *In*: WRIGHT, J. D. **International Encyclopedia of the Social & Behavioral Sciences**. 2. ed. [*S. l.*]: Elsevier, 2015. p. 360-367. DOI: https://doi.org/10.1016/B978-0-08-097086-8.92096-8. Last access on: 13 Mar 2024.

MICHAELIS, S. M.; MAURER, P.; HASPELMATH, M.; HUBER, M. (ed.). **Atlas of Pidgin and Creole language structures online**. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. Available at: http://apics-online.info. Last access on: 13 Mar 2024.

MIDGLEY, K. J.; HOLCOMB, P. J.; VAN HEUVEN, W. J.; GRAINGER, J. An electrophysiological investigation of cross-language effects of orthographic neighborhood. **Brain Research**, [*S. l.*], v. 1246, 2008. p. 123-135. DOI: https://doi.org/10.1016/j.brainres.2008.09.078. Last access on: 13 Mar 2024.

MILANKOV, V.; GOLUBOVIĆ, S.; KRSTIĆ, T.; GOLUBOVIĆ, Š. Phonological awareness as the foundation of reading acquisition in students reading in transparent orthography. **International Journal of Environmental Research and Public Health**, [*S. l.*], v. 18, n. 10, 5440, 2021. DOI: https://doi.org/10.3390/ijerph18105440. Last access on: 27 May 2024.

MISHRA, R. K.; SINGH, N. Language non-selective activation of orthography during spoken word processing in Hindi-English sequential bilinguals: an eye tracking visual world study. **Reading and Writing**: An Interdisciplinary Journal, [*S. l.*], v. 27, n. 1, p. 129-151, 2014. DOI: https://doi.org/10.1007/s11145-013-9436-5. Last access on: 13 Mar 2024.

MODESTO-SARRA, L. K. Português como língua de acolhimento: relato de uma prática pedagógica dentro de uma perspectiva intercultural. **Working Papers em Linguística**, Florianópolis, v. 23, n. 2, p. 190-210, 2022. DOI: https://doi.org/10.5007/1984-8420.2022.e83786. Last access on: 13 Mar 2024.

MONTRUL, S. Heritage languages and heritage speakers. *In*: MONTRUL, S. **The acquisition of heritage languages**. [*S. l.*]: Cambridge University Press, 2015. p. 13-40.

MOOJEN, S.; SANTOS, R. Avaliação metafonológica: resultados de uma pesquisa. **Letras de Hoje**, Porto Alegre, v. 36, n. 3, p. 751-758, 2001. Available at: https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/14639. Last access on: 5 Nov 2023.

MORAIS, J.; KOLINSKY, R.; ALEGRIA, J.; SCLIAR-CABRAL, L. Alphabetic literacy and psychological structure. **Letras de Hoje**, Porto Alegre, v. 33, n. 4, p. 61-79, 1998. Available at: https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/15117. Last access on: 5 Nov 2023.

MORTON, J. Interaction of information in word recognition. **Psychological Review**, [*S. l.*], v. 76, n. 2, p. 165-178, 1969. DOI: https://doi.org/10.1037/h0027366. Last access on: 13 Mar 2024.

MULDER, K.; DIJKSTRA, T.; BAAYEN, R. H. Cross-language activation of morphological relatives in cognates: the role of orthographic overlap and task-related processing. **Frontiers in Human Neuroscience**, [*S. l.*], v. 9, p. 1-18, 2015. DOI: https://doi.org/10.3389/fnhum.2015.00016. Last access on: 13 Mar 2024.

MULDER, K.; DIJKSTRA, T.; SCHREUDER, R.; BAAYEN, H. R. Effects of primary and secondary morphological family size in monolingual and bilingual word processing. **Journal of Memory and Language**, [*S. l.*], v. 72, p. 59-84, 2014. DOI: https://doi.org/10.1016/j.jml.2013.12.004. Last access on: 13 Mar 2024.

MULDER, K.; WLOCH, L.; BOVES, L.; TEN BOSCH, L.; ERNESTUS, M. Cognate status modulates the comprehension of isolated reduced forms. **Language, Cognition and Neuroscience**, [*S. l.*], v. 37, n. 5, p. 576-614, 2022. DOI: 10.1080/23273798.2021.1995611. Last access on: 13 Mar 2024.

MUNTENDAM, A.; VAN RIJSWIJK, R.; SEVERIJNEN, G.; DIJKSTRA, T. The role of stress position in bilingual auditory word recognition: cognate processing in Turkish and Dutch. **Bilingualism**: Language and Cognition, [*S. l.*], v. 25, n. 4, p. 679-690, 2022. DOI: https://doi.org/10.1017/S1366728922000037. Last access on: 13 Mar 2024.

MUTER, V.; HULME, C.; SNOWLING, M. J.; STEVENSON, J. Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. **Developmental Psychology**, [*S. l.*], v. 40, n. 5, p. 665-681, 2004. DOI: https://doi.org/10.1037/0012-1649.40.5.665. Last access on: 13 Mar 2024.

NATION, K. Decoding, orthographic learning, and the development of visual word recognition. *In*: SPIVEY, M.; MCRAE, K.; JOANISSE, M. (ed.). **The Cambridge handbook of Psycholinguistics**. [*S. l.*]: Cambridge University Press, 2012. p. 204-217.

NIPPOLD, M. A. Language development during the adolescent years: aspects of pragmatics, syntax, and semantics. **Topics in Language Disorders**, [*S. l.*], v. 20, n. 2, p. 15-28, 2000.

NOBLE, K. G.; MCCANDLISS, B. D.; FARAH, M. J. Socioeconomic gradients predict individual differences in neurocognitive abilities. **Developmental Science**, [*S. l.*], v. 10, n. 4, p. 464-480, 2007. DOI: https://doi.org/10.1111/j.1467-7687.2007.00600.x. Last access on: 13 Mar 2024.

OBMIGRA – OBSERVATÓRIO DAS MIGRAÇÕES INTERNACIONAIS. DataMigra. **OBMigra**, Brasília, [20-?]. Available at: https://datamigra.mj.gov.br/#/public. Last access on: 13 Mar 2024.

OLIVEIRA, A. M.; GERMANO, G. D.; CAPELLINI, S. A. Desempenho de escolares em provas de processo de identificação de letras e do processo léxico. **CEFAC**, São Paulo, v. 18, n. 5, p. 1121-1132, 2016. DOI: https://doi.org/10.1590/1982-0216201618523315. Last access on: 13 Mar 2024.

OTWINOWSKA, A.; FORYŚ-NOGALA, M.; KOBOSKO, W.; SZEWCZYK, J. Learning orthographic cognates and non-cognates in the classroom: does awareness of cross-linguistic similarity matter? **Language Learning**, [*S. l.*], v. 70, n. 3, p. 685-731, 2020. DOI: https://doi.org/10.1111/lang.12390. Last access on: 13 Mar 2024.

PALHOÇA é a cidade mais procurada por imigrantes haitianos na Grande Florianópolis. **NSC**, [*S. l.*], 15 Dec 2015. Available at: https://www.nsctotal.com.br/noticias/palhoca-e-a-cidade-mais-procurada-por-imigrantes-haitianos-na-grande-florianopolis#:~:text=A%20Grande%20Florian%C3%B3polis%20%C3%A9%20uma,e%20proximidade%20%C3%A0%20capital%20catarinense. Last access on: 13 Mar 2024.

PAULA, G. R.; MOTA, H. B.; KESKE-SOARES, M. A terapia em consciência fonológica no processo de alfabetização. **Pró-Fono Revista de Atualização Científica**, [*S. l.*], v. 17, n. 2, p. 175-184, 2005. DOI: https://doi.org/10.1590/S0104-56872005000200006. Last access on: 13 Mar 2024.

PEETERS, D.; DIJKSTRA, T.; GRAINGER, J. The representation and processing of identical cognates by late bilinguals: RT and ERP effects. **Journal of Memory and Language**, [*S. l.*], v. 68, n. 4, p. 315-332, 2013. DOI: https://doi.org/10.1016/j.jml.2012.12.003. Last access on: 13 Mar 2024.

PEREA, M. Neighborhood effects in visual word recognition and reading. *In*: POLLATSEK, A.; TREIMAN, R. (ed.). **The Oxford handbook of reading**. New York: Oxford University Press, 2015. p. 76-87.

PEREIRA, T. C. A. S.; COSTA, D. A. Línguas em contato e a formação de redes sociais de imigrantes haitianos no Rio de Janeiro. **Veredas**: Revista de Estudos Linguísticos, [*S. l.*], v. 19, n. 1, p. 92-105, 2015.

PERFETTI, C. Reading ability: lexical quality to comprehension. **Scientific Studies of Reading**, [*S. l.*], v. 11, n. 4, p. 357-383, 2007. DOI: https://doi.org/10.1080/10888430701530730. Last access on: 13 Mar 2024.

PERFETTI, C. A.; HART, L. The lexical bases of comprehension skill. *In*: GORFIEN, D. S. (ed.). **On the consequences of meaning selection**: perspectives on resolving lexical ambiguity. Washington, DC: American Psychological Association, 2001. p. 67-86. DOI: https://doi.org/10.1037/10459-004.

PICCOLO, L. R.; FALCETO, O. G.; FERNANDES, C. L.; LEVANDOWSKI, D. C.; GRASSI-OLIVEIRA, R.; SALLES, J. F. Variáveis psicossociais e desempenho em leitura de crianças de baixo nível socioeconômico. **Psicologia**: Teoria e Pesquisa, [*S. l.*], v. 28, p. 389-98, 2013. DOI: https://doi.org/10.1590/S0102-37722012000400004. Last access on: 13 Mar 2024.

PIPER, B.; BULAT, J.; JOHNSTON, A. Reading skill transfer across languages: outcomes from longitudinal bilingual randomized control trials in Kenya and Haiti. *In*: SREE SPRING 2015 CONFERENCE, 2015, [*S. l.*]. **Proceedings** […]. [*S. l.*]: SREE, 2015. p. 1-6. Available at: https://www.haiti-now.org/wp-content/uploads/2019/03/Reading-Skill-Transfer-Across-

Languages-Outcomes-from-Longitudinal-Bilingual-Randomized-Control-Trials-in-Kenya-and-Haiti.pdf. Last access on: 13 Mar 2024.

POARCH, G. J.; BIALYSTOK, E. Bilingualism as a model for multitasking. **Developmental Review**, [*S. l.*], v. 35, p. 113-124, 2015. DOI: https://doi.org/10.1016/j.dr.2014.12.003. Last access on: 13 Mar 2024.

POARCH, G. J.; VAN HELL, J. G. Cross-language activation in same-script and different-script trilinguals. **International Journal of Bilingualism**, [*S. l.*], v. 18, n. 6, p. 693-716, 2014. DOI: https://doi.org/10.1177/1367006912472262. Last access on: 13 Mar 2024.

POORT, E. D.; RODD, J. M. Towards a distributed connectionist account of cognates and interlingual homographs: evidence from semantic relatedness tasks. **PeerJ**, [*S. l.*], v. 7, e6725, 2019. DOI: https://doi.org/10.7717/peerj.6725. Last access on: 13 Mar 2024.

POPLACK, S. Sometimes I'll start a sentence in Spanish y termino en Español: toward a typology of code-switching. **Linguistics**, [*S. l.*], v. 18, n. 7-8, p. 581-618, 1980. DOI: https://doi.org/10.1515/ling.1980.18.7-8.581. Last access on: 13 Mar 2024.

POSIT TEAM. **RStudio**: Integrated Development Environment for R. Posit Software, PBC, Boston, MA, 2023.

POTTER, R. K.; KOPP, G. A.; GREEN, H. C. **Visible speech**. Naperville: D. van Nostrand Company, 1947.

PRATHEEBA, N.; KRASHEN, S. Self-reported reading as a predictor of vocabulary knowledge. **Perceptual and Motor Skills**, [*S. l.*], v. 117, n. 2, p. 442-448, 2013. DOI: https://doi.org/10.2466/23.PMS.117x24z5. Last access on: 13 Mar 2024.

PRESTON, M. S.; LAMBERT, W. E. Interlingual interference in a bilingual version of the Stroop color-word task. **Journal of Verbal Learning and Verbal Behavior**, [*S. l.*], v. 8, n. 2, p. 295-301, 1969. DOI: https://doi.org/10.1016/S0022-5371(69)80079-4. Last access on: 13 Mar 2024.

PREUSS, E. O. A funcionalidade do princípio interativo na produção de fala bilíngue. **Letras de Hoje**, Porto Alegre, v. 47, n. 1, p. 76-83, 2012. Available at: https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/9636. Last access on: 13 Mar 2024.

PREVOO, M. J. L.; MALDA, M.; MESMAN, J.; VAN IJZENDOORN, M. H. Within- and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background: a meta-analytical study. **Review of Educational Research**, [*S. l.*], v. 86, n. 1, p. 237-276, 2016. DOI: https://doi.org/10.3102/0034654315584685. Last access on: 13 Mar 2024.

PUGLISI, M.; SALLES, J. O impacto do ambiente familiar sobre o desenvolvimento cognitivo e linguístico infantil. *In*: LENT, R.; BUCHWEITZ; A.; MOTA, M. (org.). **Ciência para educação**: uma ponte entre dois mundos. São Paulo: Atheneu, 2018. v. 1, p. 221-236.

R CORE TEAM. R: a language and environment for statistical computing. **R Foundation for Statistical Computing**, Viena, 2023. Available at: https://www.R-project.org/. Last access on: 13 Mar 2024.

RAUDSZUS, H.; SEGERS, E.; VERHOEVEN, L. Situation model building ability uniquely predicts first and second language reading comprehension. **Journal of Neurolinguistics**, [*S. l.*], v. 50, p. 106-119, 2019. DOI: https://doi.org/10.1016/j.jneuroling.2018.11.003. Last access on: 13 Mar 2024.

RAYNER, K.; SCHOTTER, E.; MASSON, M.; POTTER, M. C.; TREIMAN, R. So much to read, so little time: how do we read, and can speed reading help? **Psychological Science in the Public Interest**, [*S. l.*], v. 17, p. 4-34, 2016. DOI: 10.1177/1529100615623267. Last access on: 29 May 2024.

REPPEN, R.; SIMPSON-VLACH, R. Corpus linguistics. *In*: SCHMITT, N.; RODGERS, M. P. H. (ed.). **An introduction to Applied Linguistics**. 3. ed. Routledge: London, 2019. p. 91-107.

REZAEI, A.; MOUSANEZHAD JEDDI, E. The contributions of attentional control components, phonological awareness, and working memory to reading ability. **Journal of Psycholinguistic Research**, [*S. l.*], v. 49, n. 1, p. 31-40, 2020. DOI: https://doi.org/10.1007/s10936-019-09669-4. Last access on: 13 Mar 2024.

RIGATTI, P. C.; CUI, X.; PUGH, K.; MOTA, M. B.; BUCHWEITZ, A. Brain foundations for learning to read. *In*: VERHOEVEN, L.; NAG, S.; PERFETTI, C.; PUGH, K. (ed.). **Global variation in literacy development**. [*S. l.*]: Cambridge University Press, 2023. p. 265-291.

ROCHA, N. A.; DA SILVEIRA GILENO, R. S. Portuguese as a welcoming language: breaking linguistic and cultural boundaries. *In*: DEL RÉ, A.; FALASCA, P.; NOACK NAPOLES, J. (ed.). **From discriminating to discrimination**. São Paulo: Springer, 2022. p. 93-102. DOI: https://doi.org/10.1007/978-3-031-13544-6_9. Last access on: 13 Mar 2024.

RODRIGUES, C. V. "Sou um corpo estranho no conjunto": narrativas de um estudante negro migrante em uma universidade brasileira. **Trabalhos em Linguística Aplicada**, Campinas, v. 60, n. 1, p. 114-125, 2021. DOI: https://doi.org/10.1590/01031813825511220201031. Last access on: 13 Mar 2024.

ROMBERG, A. R.; SAFFRAN, J. R. Statistical learning and language acquisition. **Wiley Interdisciplinary Reviews**: Cognitive Science, [*S. l.*], v. 1, n. 6, p. 906-914, 2010. DOI: https://doi.org/10.1002/wcs.78. Last access on: 29 May 2014.

RUDELL, A. P.; HU, B. Effects of long-time reading experience on reaction time and the recognition potential. **International Journal of Psychophysiology**, [*S. l.*], v. 76, p. 158-168, 2010. DOI: https://doi.org/10.1016/j.ijpsycho.2010.03.006. Last access on: 13 Mar 2024.

SÁ, P. R. C. de; SILVA, F. R. Desafios à inclusão dos imigrantes haitianos na sociedade brasileira. *In*: SEMINÁRIO MIGRAÇÕES INTERNACIONAIS, REFÚGIO E POLÍTICAS, MEMORIAL DA AMÉRICA LATINA, 12 Apr 2016, São Paulo. **Anais** [...]. São Paulo: Unicamp, 2016. Available at: https://www.nepo.unicamp.br/publicacoes/anais-migracoes/arquivos/27_PRCS.pdf. Last access on: 13 Mar 2024.

SAER, D. J. The effect of bilingualism on intelligence. **British Journal of Psychology**, [*S. l.*], v. 14, n. 1, 1923. DOI: https://doi.org/10.1111/j.2044-8295.1923.tb00110.x. Last access on: 13 Mar 2024.

SAIEGH-HADDAD, E. What is phonological awareness in L2? **Journal of Neurolinguistics**, [*S. l.*], v. 50, p. 17-27, 2019. DOI: https://doi.org/10.1016/j.jneuroling.2017.11.001. Last access on: 13 Mar 2024.

SAMUEL, A. G.; SUMNER, M. Current directions in research in spoken word recognition. *In*: SPIVEY, M.; MCRAE, K.; JOANISSE, M. (ed.). **The Cambridge handbook of Psycholinguistics**. Cambridge: Cambridge University Press, 2012. p. 61-75.

SÁNCHEZ-CASAS, R.: GARCÍA-ALBEA, J. E. The representation of cognate and noncognate words in bilingual memory: can cognate status be characterized as a special kind of morphological relation? *In*: KROLL, J. F.; DE GROOT, A. M. B. (ed.). **Handbook of bilingualism**: Psycholinguistic approaches. Oxford: Oxford University Press, 2005. p. 226-250.

SÁNCHEZ-CASAS, R.; ALMAGRO, Y. Efectos de priming entre lenguas utilizando primes enmascarados y no enmascarados y diferente asincronía estimular. *In*: SIMPOSIUM DE PSICOLINGUÍSTICA, 4., 1999, Madrid, Spain. [**Presented papers**]. Madrid, 1999.

SÁNCHEZ-CASAS, R.; GARCÍA-ALBEA, J. E.; IGOA, J. M. Can cognate words be characterized as a kind of morphological relations? *In*: DORÉ, F. Y. (ed.). **Abstracts of the XXVII International Congress of Psychology**. Hove, U.K.: Psychology Press, 2000. p. 405-406.

SÁNCHEZ-CASAS, R.; SUÁREZ-BURATTI, B.; IGOA, J. M. Are bilingual lexical representations interconnected? *In*: CONFERENCE OF THE EUROPEAN SOCIETY FOR COGNITIVE PSYCHOLOGY, 50., 1992, Paris, France. [**Presented papers**]. Paris, 1992.

SANTA CATARINA (Estado). SC terá Centro de Referência e Acolhimento dos Imigrantes. **Secretaria da Assistência Social, Mulher e Família**, Florianópolis, 11 Jan 2016. Available at: https://www.sas.sc.gov.br/index.php/noticias/215-sc-tera-centro-de-referencia-e-acolhimento-dos-imigrantes. Last access on: 13 Mar 2024.

SANTOS, M. T. M. dos; BEFI-LOPES, D. M. Vocabulário, consciência fonológica e nomeação rápida: contribuições para a ortografia e elaboração escrita. **Jornal da Sociedade Brasileira de Fonoaudiologia**, [*S. l.*], v. 24, n. 3, p. 269-275, 2012. DOI: https://doi.org/10.1590/S2179-64912012000300013. Last access on: 13 Mar 2024.

SANTOS-DÍAZ, I. The influence of reading on the vocabulary in mother tongue and foreign language. **Ocnos**, [*S. l.*], v. 16, n. 1, p. 79-88, 2017.

SCHRÖTER, P.; SCHROEDER, S. Orthographic processing in balanced bilingual children: cross-language evidence from cognates and false friends. **Journal of Experimental Child Psychology**, [*S. l.*], v. 141, p. 239-246, 2016. DOI: https://doi.org/10.1016/j.jecp.2015.09.005. Last access on: 13 Mar 2024.

SCHWARTZ, A. I.; KROLL, J. F. Bilingual lexical activation in sentence context. **Journal of Memory and Language**, [*S. l.*], v. 55, n. 2, p. 197-212, 2007. DOI: https://doi.org/10.1016/j.jml.2006.03.004. Last access on: 13 Mar 2024.

SEABRA, A. G. Evidências de validade e fidedignidade da Prova de Consciência Fonológica por escolha de figuras. *In*: SEABRA, A. G.; DIAS, N. M. (org.). **Avaliação neuropsicológica cognitva**: linguagem oral. São Paulo: Memnon Edições Científicas, 2013. v. 2, chapter 14. (*E-book*).

SEIDENBERG, M. S.; MCCLELLAND, J. L. A distributed, developmental model of word recognition and naming. **Psychological Review**, [*S. l.*], v. 96, n. 4, p. 523-568, 1989. DOI: https://doi.org/10.1037/0033-295X.96.4.523. Last access on: 13 Mar 2024.

SELFRIDGE, O. **Pandemonium**: a paradigm for learning, proceedings of symposium on the mechanization of thought processes. Teddington: National Physics Laboratory, 1959.

SERRATRICE, L. Becoming bilingual in early childhood. *In*: DE HOUWER, A.; ORTEGA L. (org.). **The Cambridge handbook of bilingualism**. New York: Cambridge University Press, 2019. p. 15-35.

SIEGELMAN, N.; RUECKL, J. G.; STEACY, L. M.; FROST, S. J.; VAN DEN BUNT, M.; ZEVIN, J. D.; SEIDENBERG, M. S.; PUGH, K. R.; COMPTON, D. L.; MORRIS, R. D. Individual differences in learning the regularities between orthography, phonology and semantics predict early reading skills. **Journal of Memory and Language**, [*S. l.*], v. 114, p. 104-145, 2020. DOI: https://doi.org/10.1016/j.jml.2020.104145. Last access on: 13 Mar 2024.

SMITH, F. Bilingualism and mental development. **British Journal of Psychology**, [*S. l.*], v. 13, n. 3, p. 271-282, 1923. DOI: https://doi.org/10.1111/j.2044-8295.1923.tb00101.x. Last access on: 13 Mar 2024.

SOARES, A. P.; OLIVEIRA, H. M.; COMESAÑA, M.; COSTA, A. S. Lexico-syntactic interactions in the resolution of relative clause ambiguities in a second language (L2): the role of cognate status and L2 proficiency. **Psicológica**, [*S. l.*], v. 39, n. 2, p. 164-197, 2018. DOI: https://doi.org/10.2478/psicolj-2018-0008. Last access on: 13 Mar 2024.

SOARES, A. P.; OLIVEIRA, H.; FERREIRA, M.; COMESAÑA, M.; MACEDO, A. F.; FERRÉ, P.; ACUÑA-FARIÑA, C.; HERNÁNDEZ-CABRERA, J.; FRAGA, I. Lexico-syntactic interactions during the processing of temporally ambiguous L2 relative clauses: an eye-tracking study with intermediate and advanced Portuguese-English bilinguals. **PLoS One**, [*S. l.*], v. 14, n. 5, e0216779, 2019. DOI: https://doi.org/10.1371/journal.pone.0216779. Last access on: 13 Mar 2024.

SOUZA, R. Os bilíngues interessam aos linguistas? Uma resposta da psicolinguística. *In*: OLIVEIRA, C. S. F.; SÁ, T. M. M. (org.). **Psicolinguística em Minas Gerais**. Contagem: CEFET-MG, 2020. p. 167-190.

SPARKS, R. L. Individual differences in L2 learning and long-term L1-L2 relationships. **Language Learning**, [*S. l.*], v. 62, n. suppl. 2, p. 5-27, 2012. DOI: https://doi.org/10.1111/j.1467-9922.2012.00704.x. Last access on: 13 Mar 2024.

SPIVEY, M. J.; MARIAN, V. Cross talk between native and second languages: partial activation of an irrelevant lexicon. **Psychological Science**, [*S. l.*], v. 10, n. 3, p. 281-284, 1999. DOI: https://doi.org/10.1111/1467-9280.00151. Last access on: 13 Mar 2024.

SQUIRES, L. B.; OHLFEST, S. J.; SANTORO, K. E.; ROBERTS, J. L. Factors influencing cognate performance for young multilingual children's vocabulary: a research synthesis. **American Journal of Speech-Language Pathology**, [*S. l.*], v. 29, n. 4, p. 2170-2188, 2020. DOI: https://doi.org/10.1044/2020_AJSLP-19-00167. Last access on: 13 Mar 2024.

STÄGER, L.; VON JAGEMANN, H.; JOYNES, K.; RADDATZ, C. Speaking as means and end of language teaching. **Transactions and Proceedings of the Modern Language Association of America**, [*S. l.*], v. 2, n. 1, p. xviii-xxiv, 1887. DOI: https://doi.org/10.2307/456036. Last access on: 13 Mar 2024.

STORRS, C. How poverty affects the brain. **Nature**, [*S. l.*], v. 547, p. 150-152, 2017. DOI: https://doi.org/10.1038/547150a. Last access on: 13 Mar 2024.

STRIJKERS, K.; COSTA, A.; THIERRY, G. Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. **Cerebral Cortex**, [*S. l.*], v. 20, n. 4, p. 912-928, 2010. DOI: https://doi.org/10.1093/cercor/bhp153. Last access on: 13 Mar 2024.

SUN, H.; BORNSTEIN, M. H.; ESPOSITO, G. The specificity principle in young dual language learners' English development. **Child Development**, [*S. l.*], v. 92, n. 5, p. 1752-1768, 2021. DOI: https://doi.org/10.1111/cdev.13558. Last access on: 13 Mar 2024.

SWANK, L. K.; CATTS, H. W. Phonological awareness and written word decoding. **Language, Speech, and Hearing Services in Schools**, [*S. l.*], v. 25, n. 1, p. 9-14, 1994. DOI: https://doi.org/10.1044/0161-1461.2501.09. Last access on: 13 Mar 2024.

SWINNEY, D. A. Lexical access during sentence comprehension: (re)consideration of context effects. **Journal of Verbal Learning and Verbal Behavior**, [*S. l.*], v. 18, n. 6, p. 645-659, 1979. DOI: https://doi.org/10.1016/S0022-5371(79)90355-4. Last access on: 13 Mar 2024.

SWINNEY, D. A.; ONIFER, W.; PRATHER, P.; HIRSHKOWITZ, M. Semantic facilitation across sensory modalities in the processing of individual words and sentences. **Memory & Cognition**, [*S. l.*], v. 7, n. 3, p. 159-165, 1979. DOI: https://doi.org/10.3758/BF03197534. Last access on: 13 Mar 2024.

SZUBKO-SITAREK, W. Cognate facilitation effects in trilingual word recognition. **Studies in Second Language Learning and Teaching**, [*S. l.*], v. 1, n. 2, p. 189-208, 2011. Available at: https://files.eric.ed.gov/fulltext/EJ1136468.pdf. Last access on: 13 Mar 2024.

TANAKA-ISHII, K.; TERADA, H. Word familiarity and frequency. **Studia Linguistica**, [*S. l.*], v. 65, n. 1, p. 96-116, 2011. DOI: https://doi.org/10.1111/j.1467-9582.2010.01176.x. Last access on: 13 Mar 2024.

TARGÈTE, J.; URCIOLO, R. G. **Haitian Creole-English dictionary**. Kensington: Dunwoody Press, 1993. (*E-book*).

TAYLOR, J. N.; PERFETTI, C. A. Eye movements reveal readers' lexical quality and reading experience. **Reading and Writing**: An Interdisciplinary Journal, [*S. l.*], v. 29, n. 6, p. 1069-1103, 2016. DOI: https://doi.org/10.1007/s11145-015-9616-6. Last access on: 13 Mar 2024.

THE WORLD BANK. Poverty and inequality – World Development Indicators. **The World Bank**, [*S. l.*], 2024. Available at: https://datatopics.worldbank.org/world-development-indicators/themes/poverty-and-inequality.html#:~:text=Poverty%20measured%20at%20the%20international,than%203%20percent%20by%202030. Last access on: 13 Mar 2024.

THIERRY, G.; WU, Y. J. Brain potentials reveal unconscious translation during foreign-language comprehension. **PNAS Proceedings of the National Academy of Sciences of the United States of America**, [*S. l.*], v. 104, n. 30, p. 12530-12535, 2007. DOI: https://doi.org/10.1073/pnas.0609927104. Last access on: 13 Mar 2024.

TITONE, D. A.; TIV, M. Rethinking multilingual experience through a Systems Framework of Bilingualism. **Bilingualism**: Language and Cognition, [*S. l.*], v. 26, p. 1-16, 2023. DOI: https://doi.org/10.1017/S1366728921001127. Last access on: 13 Mar 2024.

TITONE, D.; LIBBEN, M.; MERCIER, J.; WHITFORD, V.; PIVNEVA, I. Bilingual lexical access during L1 sentence reading: the effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. **Journal of Experimental Psychology**: Learning, Memory, and Cognition, [*S. l.*], v. 37, n. 6, p. 1412-1431, 2011. DOI: https://doi.org/10.1037/a0024492. Last access on: 13 Mar 2024.

TOASSI, P. F. P.; MOTA, M. B.; TEIXEIRA, E. N. O efeito de palavras cognatas no acesso lexical do inglês como terceira língua. **Cadernos de Tradução**, Florianópolis, v. 40, n. esp. 2, p. 74-96, 2020. DOI: https://doi.org/10.5007/2175-7968.2020v40nesp2p74. Last access on: 13 Mar 2024.

TOFFOLI, M. B.; LAMPRECHT, R. R. A estimulação de habilidades auditivo-verbais de crianças pré-silábicas: contribuições para o desenvolvimento da consciência fonológica. **Letras de Hoje**, [*S. l.*], v. 43, n. 3, p. 89-97, 2008. Available at: https://revistaseletronicas.pucrs.br/index.php/fale/article/view/5614. Last access on: 13 Mar 2024.

TORGESEN, J. K.; WAGNER, R. K.; RASHOTTE, C. A.; ROSE, E.; LINDAMOOD, P.; CONWAY, T.; GARVAN, C. Preventing reading failure in young children with phonological processing disabilities: group and individual responses to instruction. **Journal of Educational Psychology**, [*S. l.*], v. 91, p. 579-593, 1999. DOI: https://doi.org/doi.org/10.1037/0022-0663.91.4.579. Last access on: 29 May 2024.

TUCKER, B. V.; BRENNER, D.; DANIELSON, D. K.; KELLEY, M. C.; NENADIĆ, F.; SIMS, M. The Massive Auditory Lexical Decision (MALD) database. **Behavior Research Methods**, [*S. l.*], v. 51, p. 1187-1204, 2019. DOI: https://doi.org/10.3758/s13428-018-1056-1. Last access on: 13 Mar 2024.

TYTUS, A. E. Active and dormant languages in the multilingual mental lexicon. **International Journal of Multilingualism**, [*S. l.*], v. 16, n. 3, 2019, p. 357-374. DOI: https://doi.org/’0.1080/14790718.2018.1502295. Last access on: 13 Mar 2024.

UN – UNITED NATIONS. MINUSTAH fact sheet. **United Nations Peacekeeping**, [*S. l.*], [20--]. Available at: https://peacekeeping.un.org/en/mission/minustah. Last access on: 13 Mar 2024.

UNDP – UNITED NATIONS DEVELOPMENT PROGRAMME. Human Development Index (HDI). **UNDP**, [*S. l.*], 2023. Available at: https://hdr.undp.org/data-center/human-development-index#/indicies/HDI. Last access on: 13 Mar 2024.

UNHCR – UNITED NATIONS HIGH COMMISSIONER FOR REFUGEES. Com apoio do ACNUR, governo lança plano de ação para população haitiana no Brasil. **United Nations High Commissioner for Refugees**, [*S. l.*], 20 Jun 2023. Available at: https://www.acnur.org/portugues/2023/06/20/com-apoio-do-acnur-governo-lanca-plano-de-acao-para-populacao-haitiana-no-brasil/#:~:text=Estima%2Dse%20que%2C%20atualmente%2C,integra%C3%A7%C3%A3o%20local%20a%20serem%20enfrentados. Last access on: 13 Mar 2024.

UNIVERSITY OF MARYLAND. Maryland Language Science Center. **LangScape** – Statistics, College Park, 2018. Available at: http://langscape.umd.edu/stats.php. Last access on: 13 Mar 2024.

VAN ASSCHE, E.; BRYSBAERT, M.; DUYCK, W. Bilingual lexical access. *In*: HEREDIA, R. R.; CIESLICKA, A. B. (ed.). **Bilingual lexical ambiguity resolution**. [*S. l.*]: Cambridge University Press, 2020. p. 42-67. DOI: https://doi.org/10.1017/9781316535967. Last access on: 13 Mar 2024.

VAN ASSCHE, E.; DRIEGHE, D.; DUYCK, W.; WELVAERT, M.; HARTSUIKER, R. J. The influence of semantic constraints on bilingual word recognition during sentence reading. **Journal of Memory and Language**, [*S. l.*], v. 64, n. 1, p. 88-107, 2011. DOI: https://doi.org/10.1016/j.jml.2010.08.006. Last access on: 13 Mar 2024.

VAN ASSCHE, E.; DUYCK, W.; HARTSUIKER, R. J.; DIEPENDAELE, K. Does bilingualism change native-language reading? Cognate effects in a sentence context. **Psychological Science**, [*S. l.*], v. 20, n. 8, p. 923-927, 2009. Available at: http://www.jstor.org/stable/40575121. Last access on: 13 Mar 2024.

VAN DIJK, C.; DIJKSTRA, T.; UNSWORTH, S. Cross-linguistic influence during online sentence processing in bilingual children. **Bilingualism**: Language and Cognition, [*S. l.*], v. 25, n. 4, p. 691-704, 2022. DOI: https://doi.org/10.1017/S1366728922000050. Last access on: 13 Mar 2024.

VAN ORDEN, G. C. A ROWS is a ROSE: spelling, sound, and reading. **Memory & Cognition**, [*S. l.*], v. 15, n. 3, p. 181-198, 1987. DOI: https://doi.org/10.3758/BF03197716. Last access on: 13 Mar 2024.

VASQUEZ CAMPOS, I. **Implementing explicit instruction of cognate vocabulary with migrant English language learners**. 2023. Dissertation (Doctor of Philosophy) – College of

Health, Education and Social Transformation, New Mexico State University, New Mexico, 2023.

VERHOEVEN, L. Components in early second language reading and spelling. **Scientific Studies of Reading**, [*S. l.*], v. 4, n. 4, p. 313-330, 2000.

VERHOEVEN, L. Early bilingualism, language transfer, and phonological awareness. **Applied Psycholinguistics**, [*S. l.*], v. 28, p. 425-439, 2007.

VERHOEVEN, L.; VAN LEEUWE, J. The simple view of second language reading throughout the primary grades. **Reading and Writing**, [*S. l.*], v. 25, p. 1805-1818, 2012.

VERHOEVEN, L.; VOETEN, M.; VERMEER, A. Beyond the simple view of early first and second language reading: the impact of lexical quality. **Journal of Neurolinguistics**, [*S. l.*], v. 50, p. 28-36, 2019.

VON GREBMER ZU WOLFSTHURN, S.; PABLOS, L.; SCHILLER, N. O. Noun-phrase production as a window to language selection: an ERP study. **Neuropsychologia**, [*S. l.*], v. 162, e108055, 2021. DOI: https://doi.org/10.1016/j.neuropsychologia.2021.108055. Last access on: 13 Mar 2024.

WAGNER, R. K.; PIASTA, S. B.; TORGESEN, J. K. Learning to read. *In*: TRAXLER, M. J.; GERNSBACHER, M. A. (ed.). **Handbook of Psycholinguistics**. 2. ed. London: Elsevier, 2006. p. 1111-1143.

WAGNER, R. K.; TORGESEN, J. K.; RASHOTTE, C. A.; HECHT, S. A.; BARKER, T. A.; BURGESS, S. R.; DONAHUE, J.; GARON, T. Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. **Developmental Psychology**, [*S. l.*], v. 33, n. 3, p. 468-479, 1997. DOI: https://doi.org/10.1037//0012-1649.33.3.468. Last access on: 13 Mar 2024.

WALSH, D. D. Bilingualism and bilingual education: a guest editorial. **Foreign Language Annals**, [*S. l.*], v. 2, n. 3, p. 298-303, 1969. DOI: https://doi.org/10.1111/j.1944-9720.1969.tb00306.x. Last access on: 13 Mar 2024.

WANG, A.; YEON, J.; ZHOU, W.; SHU, H.; YAN, M. Cross-language parafoveal semantic processing: evidence from Korean-Chinese bilinguals. **Psychonomic Bulletin & Review**, [*S. l.*], v. 23, n. 1, p. 285-290, 2016. DOI: https://doi.org/10.3758/s13423-015-0876-6. Last access on: 13 Mar 2024.

WILLIAMS, J. N. The relationship between word meanings in the first and second language: evidence for a common, but restricted, semantic code. **European Journal of Cognitive Psychology**, [*S. l.*], v. 6, p. 195-220, 1994. DOI: https://doi.org/10.1080/09541449408520143. Last access on: 13 Mar 2024.

WINTHER, I. E.; MATUSEVYCH, Y.; PICKERING, M. J. Can learning explain cognate effects in bilingual comprehension and production? *In*: SANTESTEBAN, M.; DUÑABEITIA, J. A.; BAUS, C. (ed.). Bilingualism through the prism of Psycholinguistics: in honour of Albert Costa. **Bilingual Processing and Acquisition**, [*S. l.*], v. 17, p. 104-129, 2023. DOI: https://doi.org/10.1075/bpa.17.03win. Last access on: 13 Mar 2024.

WOOLPERT, D. The impact of Spanish on the English lexical processing of dual-language learners. **TESOL Quarterly**, [*S. l.*], v. 53, n. 1, p. 221-232, 2018. DOI: https://doi.org/10.1002/tesq.444. Last access on: 13 Mar 2024.

WOUMANS, E.; CLAUWS, R.; DUYCK, W. Hands down: cognate effects persist during written word production. **Frontiers in Psychology**, [*S. l.*], v. 12, e647362, 2021. DOI: https://doi.org/10.3389/fpsyg.2021.647362. Last access on: 13 Mar 2024.

XIONG, K.; VERDONSCHOT, R. G.; TAMAOKA, K. The time course of brain activity in reading identical cognates: an ERP study of Chinese-Japanese bilinguals. **Journal of Neurolinguistics**, [*S. l.*], v. 55, e100911, 2020. DOI: https://doi.org/10.1016/j.jneuroling.2020.100911. Last access on: 13 Mar 2024.

YAP, M. J.; BALOTA, D. A. Visual word recognition. *In*: POLLATSEK, A.; TREIMAN, R. **The Oxford handbook of reading**. New York: Oxford University Press, 2015. p. 26-43.

YATES, M.; LOCKER JR, L.; SIMPSON, G. B. The influence of phonological neighborhood on visual word perception. **Psychonomic Bulletin & Review**, [*S. l.*], v. 11, n. 3, p. 452-457, 2004. DOI: https://doi.org/10.3758/bf03196594. Last access on: 13 Mar 2024.

ZHANG, J.; WU, C.; ZHOU, T.; MENG, Y. Cognate facilitation priming effect is modulated by writing system: evidence from Chinese-English bilinguals. **International Journal of Bilingualism**, [*S. l.*], v. 23, n. 2, p. 553-566, 2019. DOI: https://doi.org/10.1177/1367006917749062. Last access on: 13 Mar 2024.

ZIMMERMANN, N.; DELAERE, F. J.; FONSECA, R. P. **Tarefas para avaliação neuropsicológica (3)**: avaliação de memória episódica, percepção, linguagem e componentes executivos para adultos. São Paulo: Memnon, 2019.

ZUNINI, R. A. L.; BAART, M.; SAMUEL, A. G.; ARMSTRONG, B. C. Lexico-semantic access and audiovisual integration in the aging brain: insights from mixed-effects regression analyses of event-related potentials. **Neuropsychologia**, [*S. l.*], v. 165, e108107, 2022. DOI: https://doi.org/10.1016/j.neuropsychologia.2021.108107. Last access on: 13 Mar 2024.

APPENDIX A – Selected questions from the LHQ 3.0 in BP

ID ______ DEVOLVER PARA ESCOLA ____/____/2022

Questionário sobre histórico linguístico

As perguntas deste questionário são sobre o histórico linguístico do(a) seu(sua) filho(a). Por favor, responda as perguntas em relação a ele(ela).

Idade do(a) filho(a):

Sexo: ( ) Masculino ( ) Feminino

( ) Não quero responder

Escolaridade do(a) filho(a):

( ) Ensino infantil ( ) Ensino fundamental

( ) Ensino médio

Escolaridade da mãe:

( ) Ensino infantil ( ) Ensino fundamental

( ) Ensino médio ( ) Graduação ( ) Mestrado

( ) Doutorado ( ) Outros

Escolaridade do pai:

( ) Ensino infantil ( ) Ensino fundamental

( ) Ensino médio ( ) Graduação ( ) Mestrado

( ) Doutorado ( ) Outros

Idade em que o(a) filho(a) começou a aprender cada habilidade linguística em cada língua que fala:

Língua	Escutar	Falar	Ler	Escrever

Total de anos que levou para aprender cada habilidade linguística em cada língua que fala:

Língua	Escutar	Falar	Ler	Escrever

País de origem ou nascimento:

País de residência:

Se o(a) filho(a) viajou ou morou em outros países por mais de 3 meses, indique a língua que ele(a) usou, por quanto tempo e com quanta frequência, de acordo com esta escala:

1 Nunca

2 Raramente

3 Ocasionalmente

4 Algumas vezes

5 Frequentemente

6 Muito frequentemente

7 Sempre

País	Duração em meses	Língua	Frequência de uso da língua

Como o(a) filho(a) aprendeu cada língua não-materna que fala?

Língua	Imersão na sociedade	Sala de aula	Estudando sozinho(a)

Idade em que o(a) filho(a) começou a aprender cada língua nesses ambientes:

Língua	Em casa	Com amigos	Na escola	No comput ador	Em jogos online

---

Como você, responsável, avalia as habilidades linguísticas do(a) filho(a) nas línguas que ele(a) fala, de acordo com a escala abaixo?

1 Muito ruim   2 Ruim   3 Razoável   4 Funcional

5 Bom   6 Muito bom   7 Excelente

Língua	Escutar	Falar	Ler	Escrever

Durante quantas horas por dia o(a) filho(a) faz as seguintes atividades em cada língua que fala?

Língua	Assistir à TV	Ouvir rádio	Ler para se divertir	Ler para escola	Escrever para escola	Redes sociais

---

Durante quantas horas por dia o(a) filho(a) usa cada língua para conversar com estas pessoas?

Língua	Família	Amigos	Colegas de escola

Usando qual língua o(a) filho(a) se sente mais confortável nestas habilidades e nestes ambientes?

	Escutar	Falar	Ler	Escrever
Em casa				
Com amigos				
Na escola				

Com que frequência o(a) filho(a) usa as línguas que fala para estas atividades?

1 Nunca — 2 Raramente — 3 Ocasionalmente — 4 Algumas vezes

5 Frequentemente — 6 Muito frequentemente — 7 Sempre

Língua	Pensar	Falar consigo mesmo	Expressar raiva ou carinho	Sonhar	Fazer cálculos matemático	Lembrar telefones e endereços	Rezar/orar

Seu(sua) filho(a) tem alguma dificuldade de audição ou de visão?
( ) Sim ( ) Não

Se sim, está corrigida? ( ) Sim ( ) Não

APPENDIX B – SELECTED QUESTIONS FROM THE LHQ 3.0 IN HC

UFSC LabLing

ID ______ **REMÈT LEKÒL LA ____/____/2022**

Kesyonè sou istorik lengwistik

Kesyon yo ki nan kesyonè sa a se sou istorik lengwistik pitit ou a. Tanpri,reponn kesyon sa yo an relasyon a li menm.

Laj pitit ou a:

Sèks: ( ) Maskilen ( ) Feminen ( ) Mwen pa vle reponn

Nivo eskolarite pitit ou a:

( ) Jaden danfans ( ) Etid fondamantal ( ) Etid segondè

Eskolarite Manman an:

( ) jaden danfans ( ) Etid fondamantal ( ) Etid segondè

( ) Gradyasyon ( ) Metriz ( ) Doktora ( ) lòt

Eskolarite Papa a:

( ) jaden danfans ( ) Etid fondamantal ( ) Etid segondè

( ) Gradyasyon ( ) Metriz ( ) Doktora ( ) lòt

Nan ki laj pitit ou a kòmanse aprann chak abilite lengwistik nan chak lang li pale:

Lang	Tande	Pale	Li	Ekri

Peyi orijin oubyen peyi nesans:

Peyi de residans:

Total ane ke li pran poul aprann chak abilite lengwistik nan chak lang li pale:

Lang	Tande	Pale	Li	Ekri

Si pitit ou a te abite nan lòt peyi pou pi plis ke 3 mwa ,endike ki lang ke li te itilize,pou konbyen tan e konbyen fwa,an akò avèk eskal sa a:

1 Pa janm   2 Raman   3 Kèk okazyon   4 Kèk fwa

5 Souvan   6 Trè souvan   7 Toujou

Peyi	Dirasyon an mwa	Lang	Nonb fwa li itilize lang lan

Kòman pitit ou a aprann chak lang ki pa lang matènèl ke li pale?

Lang	Andedan sosyete a	Nan Sal de klas	Etidye pou kont li

Laj ke pitit ou a kòmanse aprann chak lang nan milye sa yo:

Lang	Lakay	Avèk zanmi	Nan lekòl	Sou òdinatè	Nan jwèt online

Koman ou menm, responsab, ou evalye abilite lengwistik pitit ou a nan lang ke li pale,an akò avèk eskal ki anba a?

1 Trè movè

2 Movè

3 akseptab

4 Fonksyonèl

5 Bon

6 Trè bon

7 Ekselan

Lang	Tande	Pale	Li	Ekri

Diran konbyen è pa jou pitit ou a fè aktivite sa yo a chak lang ke li pale?

Lang	Asiste TV	ekoute radyo	Li pou li pran plezi	Li pou lekòl	Ekri pou lekòl	Rezo sosyal yo

---

Diran konbyen è pa jou pitit ou a itilize chak lang pou li pale avèk moun sa yo?

Lang	Fanmiy	Zanmi	Kòlèg klas

Ki lang pitit ou a santi ke li pi konfòtab nan abilite sa yo e nan milye sa yo?

	Tande	Pale	Li	Ekri
Lakay				

Avèk zanmi				
Nan lekòl				

Avèk ki frekans pitit ou a itilize lang sa yo ke li pale pou aktivite sa yo?

1 2 3 4

Pa janm Raman kèk okazyon Kèk fwa

5 6 7

Souva Trè souvan Toujou

Lang	Panse	Pale avèk mwen menm	Eksprime kòlè oubyen jantiyès	Reve	Fé kalkil matematik	Sonje telefòn ak adrès	Pou lapriyè

Ou genyen kèk difikilte pou ou tande oubyen difikilte vizyèl? () Wi ()Non

Si Wi, sa corije? () Wi () Non

APPENDIX C – Consent form in BP used in Study 1

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA
LABORATÓRIO DA LINGUAGEM E PROCESSOS COGNITIVOS

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO – TCLE
baseado na Resolução 510/16 do CNS (Conselho Nacional de Saúde)

Pesquisa: **INFLUÊNCIAS DA COATIVAÇÃO DA L1 NA LEITURA EM L2 POR BILÍNGUES CRIOULO HAITIANO-PORTUGUÊS BRASILEIRO**

Prezados Pais ou Responsáveis,

Eu, Pietra Cassol Rigatti, aluna de Doutorado do Programa de Pós-Graduação em Linguística sob orientação da Prof.ª Dr.ª Mailce Borges Mota, na Universidade Federal de Santa Catarina (UFSC), venho por meio deste solicitar sua permissão para convidar seu(sua) filho(a) a colaborar com minha pesquisa.

O objetivo geral desta pesquisa é investigar a influência do crioulo haitiano como primeira língua (L1) na leitura em português brasileiro como segunda língua (L2) em crianças e adultos falantes nativos de crioulo haitiano.

Por isso, gostaríamos de convidar seu(sua) filho(a) para participar como voluntário(a) deste estudo, com seu consentimento. Seu(sua) filho(a) será solicitado/a a realizar as seguintes atividades:

1. **Duas tarefas com palavras em português, uma com palavras escritas e outra com palavras faladas:** São tarefas em que seu(sua) filho(a) vai decidir se uma série de letras escritas ou faladas é uma palavra de verdade ou não.

2. **Alguns testes de consciência fonológica em português:** Estas são atividades de curta duração adequadas à faixa etária em que seu(sua) filho(a) vai responder perguntas como "qual palavra rima com cama?", "qual dessas palavras começa com o mesmo som: areia, bota, arena ou maçã?".

3. **Teste de identificação de letras:** Trata-se de duas atividades de curta duração para verificar se seu(sua) filho(a) conhece as letras do alfabeto e se sabe seus nomes.

4. **Uma tarefa de vocabulário em crioulo haitiano:** Nesta tarefa, seu(sua) filho(a) vai ouvir algumas palavras em crioulo haitiano e vai decidir se as conhece ou não.

5. **Questionário sobre línguas utilizadas pela criança:** Por fim, os pais ou responsáveis serão solicitados a responderem pela criança a um questionário sobre quais línguas a criança utiliza em locais diferentes e com pessoas diferentes.

A sessão de participação e as tarefas serão realizadas de forma remota, por plataformas como o Formulário Google e o Zoom. A sessão de coleta de dados será realizada na escola em que seu(sua) filho(a) está matriculado(a) em uma sala disponibilizada pela escola, utilizando um computador da escola. A sessão poderá ser realizada no turno de aula em que a criança está matriculada ou no turno inverso. Caso ocorra no turno inverso, a despesa de transporte será ressarcida pelas pesquisadoras. A realização destas atividades pode causar algum desconforto, tédio, nervosismo ou cansaço físico para seu(sua) filho(a). Para evitar que as atividades sejam cansativas ou desconfortáveis, garantiremos intervalos entre os testes. A sessão de participação deve demorar em torno de 1h. Garantimos que seu(sua) filho(a) pode desistir a qualquer momento sem prejuízo de qualquer natureza para ele(ela), basta entrar em contato com as pesquisadoras.

A participação de seu(sua) filho(a) não implicará prejuízos ou divulgação de nomes ou identificação dos participantes de qualquer forma. Cada criança receberá um código de identificação, de forma a deixar a identidade do(a) seu(sua) filho(a) anônima na pesquisa. Informações como nome e idade serão registradas apenas com o intuito de controle da pesquisadora. Mesmo que não seja a vontade da pesquisadora, pode acontecer de outras pessoas terem acesso às respostas e informações pessoais dos participantes. Para evitar que isso aconteça, as atividades serão realizadas individualmente com cada criança, e as respostas delas serão usadas apenas para fins de pesquisa científica e ficarão armazenadas anonimamente em um dispositivo acessível somente por senha pessoal da pesquisadora responsável.

Há riscos leves ao participar desta pesquisa. Esses riscos são relacionados a situações de avaliação, como sensação de nervosismo, aborrecimento e/ou constrangimento. Também pode ocorrer cansaço físico, já que a sessão de coleta de dados pode durar de 45min a 1h30min e ocorrerá majoritariamente em frente a uma tela de computador tanto em formato presencial quanto remoto. Para evitar esses riscos, haverá intervalos programados para descansar e se hidratar. Lembramos também que é possível desistir a qualquer momento da pesquisa, sem qualquer prejuízo. Participar da pesquisa não oferece benefícios diretos aos participantes, apenas contribui para a ciência brasileira.

Os resultados desta pesquisa poderão ser divulgados em eventos ou publicações científicas e disponibilizados anonimamente em plataformas seguras, sem revelar nenhuma informação pessoal que possa identificar você ou seu(sua) filho(a). O(A) Sr.(ª) pode receber os resultados a qualquer momento, é só entrar em contato com as pesquisadoras. O(A) Sr.(ª) ou seu(sua) filho(a) apenas poderão receber ressarcimento caso haja eventuais despesas em decorrência da participação na pesquisa. Caso isso ocorra, será devolvido o valor que você ou seu(sua) filho(a) gastaram. Se o(a) Sr.(ª) ou seu(sua) filho(a) tiverem prejuízos por causa da pesquisa, vocês têm direito à indenização.

Nós, pesquisadoras, nos comprometemos a realizar a pesquisa de acordo com a Resolução do Conselho Nacional de Saúde no 510, de 07 de abril de 2016, que estabelece as normas éticas para as pesquisas em Ciências Humanas e Sociais. O Comitê de Ética em Pesquisas com Seres Humanos da Universidade Federal de Santa Catarina (CEPSH-UFSC) é o responsável por aprovar esta pesquisa. O CEPSH-UFSC foi criado para defender os seus direitos e garantir que eles sejam respeitados e que a pesquisa seja realizada de forma ética, assegurando todos os seus direitos e bem estar.

Informo que o(a) Sr.(ª) tem a garantia de acesso, em qualquer etapa do estudo, a qualquer esclarecimento sobre o estudo. Se tiver alguma consideração ou dúvida sobre a pesquisa, entre em contato pelo e-mail: pietracr@gmail.com; ou pelo telefone (51)99545-0025; ou também com a Prof.ª Dr.ª Mailce Borges Mota através do e-mail mailce.mota@ufsc.br. Você também pode entrar em contato com o Comitê de Ética em Pesquisa com Seres Humanos da UFSC através do e-mail cep.propesq@contato.ufsc.br, do telefone (48) 3721-6094, da página da internet https://cep.ufsc.br/contato/ ou no seguinte endereço: Rua Desembargador Vitor Lima, n° 222, 4° andar, sala 401, bairro Trindade, CEP 88040-400, Florianópolis, Santa Catarina.

Como informado acima, é garantida a liberdade da retirada de consentimento a qualquer momento e seu(sua) filho(a) pode deixar de participar do estudo, sem qualquer prejuízo ou punição. A retirada pode ser solicitada diretamente com a pesquisadora, por telefone ou e-mail.

Este documento será assinado em duas vias, uma das quais permanecerá em posse do(a) responsável, que deve guardá-la.

Por favor escolha uma das opções:

( ) Concordo voluntariamente em permitir a participação do(a) meu(minha) filho(a) na pesquisa **INFLUÊNCIAS DA COATIVAÇÃO DA L1 NA LEITURA EM L2 POR BILÍNGUES CRIOULO HAITIANO-PORTUGUÊS BRASILEIRO,** de autoria de Pietra Cassol Rigatti. Declaro que li este documento e que compreendi as informações do Termo de Consentimento Livre e Esclarecido. Eu compreendo meus direitos como responsável e os direitos do meu(minha) filho(a) e permito a participação dele(dela) neste estudo e em ceder os dados do meu(minha) filho(a) para a pesquisa. Compreendo o objetivo do estudo, a forma como ele será realizado e a possibilidade de retirar o meu consentimento a qualquer momento, antes ou durante a pesquisa, sem penalidade, prejuízo ou perda de qualquer benefício que eu possa ter adquirido.

( ) Eu não autorizo meu filho(a) a participar da pesquisa.

APPENDIX D – Consent form in HC used in Study 1

INIVÈSITE FEDERAL SANTA CATARINA
SANT DE KOMINIKASYON E EKSPRESYON
PWOGRAM APWÈ-GRADYASYON AN LENGWISTIK
LABORATWA AN LANG E PWOSESIS COGNITIF (MANTAL)

TÈM CONSANTMAN LIB E EKLÈSISMAN – TCLE
Baze nan Rezolisyon 510/16 do CNS (Konsèy Nasyonal Sante)

Rechèch: **ENFLIYANS NAN KOAKTIVASYON L1 NAN LEKTI AN L2 POU BILENG KREYÒL AYISYEN-PÒTIGÈ BREZILYEN**

Chè Paran oubyen chè Responsab,

Mwen menm, Pietra Cassol Rigatti, elèv kap fè Metriz nan Pwogram Apwè Gradyasyon an Lengwistik sou oryantasyon Prof.ª Dr.ª Mailce Borges Mota, nan Inivèsite Federal Santa Catarina (UFSC),Mwen vini pou mwen mande ou pèmisyon pou envite pitit ou a kolabore avèk rechèch mwen an.

Objektik general rechèch sa a se mennen ankèt sou enflyans kreyòl ayisyen kòm premye lang (L1) nan lekti an pòtigè brezilyen kòm dezyèm lang (L2) pou timoun e jèn ki pale kreyòl ayisyen natif natal.

Se pou sa,nou ta renmen envite pitit ou a pou li patisipe kòm volontè nan etid sa a,avèk konsantman ou. Nou pwal mande pou Pitit ou a reyalize aktivite sa yo:

1. **De (2) Devwa avék mo an pòtigè,yonn avèk mo ekri e lòt la avèk mo pale:** Se devwa ke pitit ou a pwal deside si seri de lèt ekri oubyen pale sa yo se yon vwè mo yo ye oubyen non.

2. **Kèk tès de konsyans fonolojik an pòtigè:** Sa yo se aktivite kout nan tan adekwa pa tranch daj,kote pitit ou a pwal reponn kesyon tankou “ ki mo ki rime avék kabann (cama)?", "ki mo nan mo sa yo ki kòmanse avèk menm son: sab (areia),bòt (bota),tèren (arena) ou pòm (maçã)?".

3. **Tès de identifikasyon de lèt:** Trete de (2) aktivite de kout dire pou verifye se pitit ou a konnen lèt nan alfabèt a e se li konnen non yo tou.

4. **Yon devwa vokabilè an kreyòl ayisyen:** Nan devwa sa a pitit ou a pwal tande kèk mo an kreyòl Ayisyen epi ou pwal deside si ou konnen yo oubyen non.

5. **Kesyonè sou lang ke pitit ou a itilize:** Pou fini, nou pwal mande paran yo ou responsab yo reponn pou pitit yo a yon kesyonè ki lang pitit yo a itilize nan lokal diferan e avèk moun diferan.

Sesyon patisipasyon an e devwa yo pwal realize sou fòm online,nan platafòm tankou Fòmilè Google e nan Zoom. Sesyon pran done yo pwal reyalize nan lekòl ke pitit ou a matrikile a nan yon sal ke lekòl la ap disponibilize pou sa, yap itilize òdinatè lekòl la. Sesyon an pwal reyalize nan menm lè kou ke pitit ou a matrikile a oubyen nan yon lòt lè ki kontrè ak lè pitit ou a matrikile a. An ka sa ta rive moun kap fè rechèch yo ap remèt lajan ki depanse pou transpò Timoun nan. Reyalizasyon aktivite sa yo kapab fè pitit ou a santi kèk ti malèz,annwi,nève ou bouke fizikman. Pou evite ke aktivite sa yo pa bay fatig oubyen malèz,nou garanti nou ke ap genyen entèval tan rekreyasyon nan tès yo. Sesyon patisipasyon an dwe dire 1h. Nou garanti ou pitit ou a kapab abandone a nepòt moman na li pap pèdi anyen,sèlman depi ou antre an kontak avèk moun kap fè rechèch la.

Patisipasyon pitit ou a pap pwal enplike okenn pèt oubyen pibliye non li oubyen idantifikasyon patisipan yo, sou nepòt fòm. Chak timoun pwal resevwa yon kòd idantifikasyon,nan fòm pou nou kite idantite pitit ou a anonim nan rechèch sa. Enfòmasyon tankou non ak laj pitit ou a pwal anrejistre jis pou moun kap fè rechéch la fè kontwòl. Menm si se pa volonte moun kap fè rechèch la ,sa ka rive ke lòt moun gen aksè avèk repons enfómasyon pèsonel patisipan yo. Pou evite sa pa rive ,aktivite yo pwal reyalize endividyèlman avèk chak timoun ,epi repons yo pwal itilize sélman nan fen rechèch syantifik e yo pwal mete yo nan yon dispozitif ki gen aksè sélman avèk kòd pèsonèl moun ki responsab fè rechéch la.

Gen ti risk tou piti nan patisipasyon rechèch sa a.Risk sa yo se risk ki rive nan sitiyasyon evalyasyon,tankou sansasyon nève ,kontraryete,oubyen anbarasman.Kapab genyen fatig fizikman ,paske kolesyon done yo kapab dire 45 minit a 1h30 minit e sa pwal fèt majoritèman devan yon ekran òdinatè ni prezansyèl oubyen online. Pou evite risk sa yo,pwal genyen entèval pwograme pou pran ti repo ,bwè dlo pou idrate nou . Sonje tou li posib pou ou abandone a nepòt ki moman nan rechèch la,san okenn pèt . Patisipe nan rechèch sa a pa ofri okenn benefis dirèk a patisipan yo,sèlman pou kontribye nan syans brezilyen an .

Rezilta rechèch sa a kapab ale pibliye nan evènman oubyen nan piblikasyon syantifik e disponibilize nan platfòm ki an sekirite,san revele okenn enfòmasyon pèsonèl ki kapab idantifye ou menm oubyen pitit ou a . Ou kapab resevwa rezilta yo a nepòt kèl moman,se sèlaman antre an kontak avèk moun kap fè rechèch yo. Ou menm oubyen pitit ou a kapab resevwa ranbousman an ka si gen yon evantyèl depans a koz patisipasyon ou nan rechèch sa. An ka ke sa rive vre ,n´ap remèt ou oubyen pitit ou a valè ki depanse a . Si ou menm oubyen pitit ou a fè pèt akoz rechèch sa,nou genyen dwa pou yo dedomaje nou.

Nou menm, moun kap fè rechèch, nou pwomèt reyalize rechèch la an akò avèk Rezolisyon Konsèy nasyonal sante nan 510, 07 avril 2016, ki etabli nòm ak etik pou rechèch an Syans Imèn e Sosyal . Komite Etik an Rechèch avèk Èt Imen nan Inivèsite Federal Santa Catarina (CEPSH-UFSC) se responsab pou apwouve rechèch sa. CEPSH-UFSC te kreye pou defann dwa ou yo e garanti ke yo respekte e tou pou rechèch yo reyalize nan fòm etik,an sekirite, a tout dwa yo e byennèt sosyal.

Mwen ap enfòme ou ke ou genyen garanti ak aksè ,nan nepòt etap etid sa ,nepòt eklèsisman sou etid sa.Si ou genyen kèk konsiderasyon oubyen dout sou rechèch la ,antre an kontak sou e-mail: pietracr@gmail.com; oubyen sou telefòn (51)99545-0025; oubyen tou avèk Prof.ª Dr.ª Mailce Borges Mota atravè e-mail: mailce.mota@ufsc.br. Ou kapab antre na kontak tou avèk Komite Etik an rechèch avèk Èt Imen nan UFSC atravè e-mail cep.propesq@contato.ufsc.br, nan telefòn (48) 3721-6094, nan paj entènèt https://cep.ufsc.br/contato/ oubyen nan adrès sa: Rua Desembargador Vitor Lima, n° 222, 4° etaj, sal 401, bairro Trindade, CEP 88040-400, Florianópolis, Santa Catarina.

Menm jan sa enfòme anlè a,li garanti ou libète pou ou retire konsantman ou a nepòt ki moman e pitit ou a kapab abandone e pa patisipe nan etid la ankò san okenn pèt ou pinisyon. Demand sa a kapab fèt dirèk avèk moun kap fè rechèch la ,nan telefòn oubyen sou e-mail. Dokiman sa a pwal siyen an 2 kopi,yonn ladan yo pwal rete pèmanan nan men responsab la ,ki genyen pou dwa sere li.

Tanpri chwazi yon opsyon:

( ) Mwen dakò volontèman pèmèt ke pitit mwen an patisipe nan rechèch sa **ENFLIYANS NAN KOATIVASYON L1 NAN LEKTI AN L2 POU BILENG KREYÒL AYISYEN-PÒTIGÈ BREZILYEN,** nan patènite Pietra Cassol Rigatti. Deklare ke mwen li dokiman sa e mwen konprann enfòmasyon yo tèm konsantman lib e eklèsisman. Mwen konprann dwa mwen yo kóm responsab e dwa pitit mwen an e pèmèt patisipasyon li nan etid rechèch sa e bay done pitit mwen an nan rechèch sa.Mwen konprann objektif etid la ,sou fòm li pwal reyalize e posibilite retire konsantman mwen a nepòt ki moman ,avan oubyen pandan rechèch la ,san penalite ,san pèdi okenn benefis ke mwen kapab jwenn.

( ) Mwen pa otorize pitit mwen an patisipe nan rechèch la.

APPENDIX E – Willingness form in BP used in Study 1

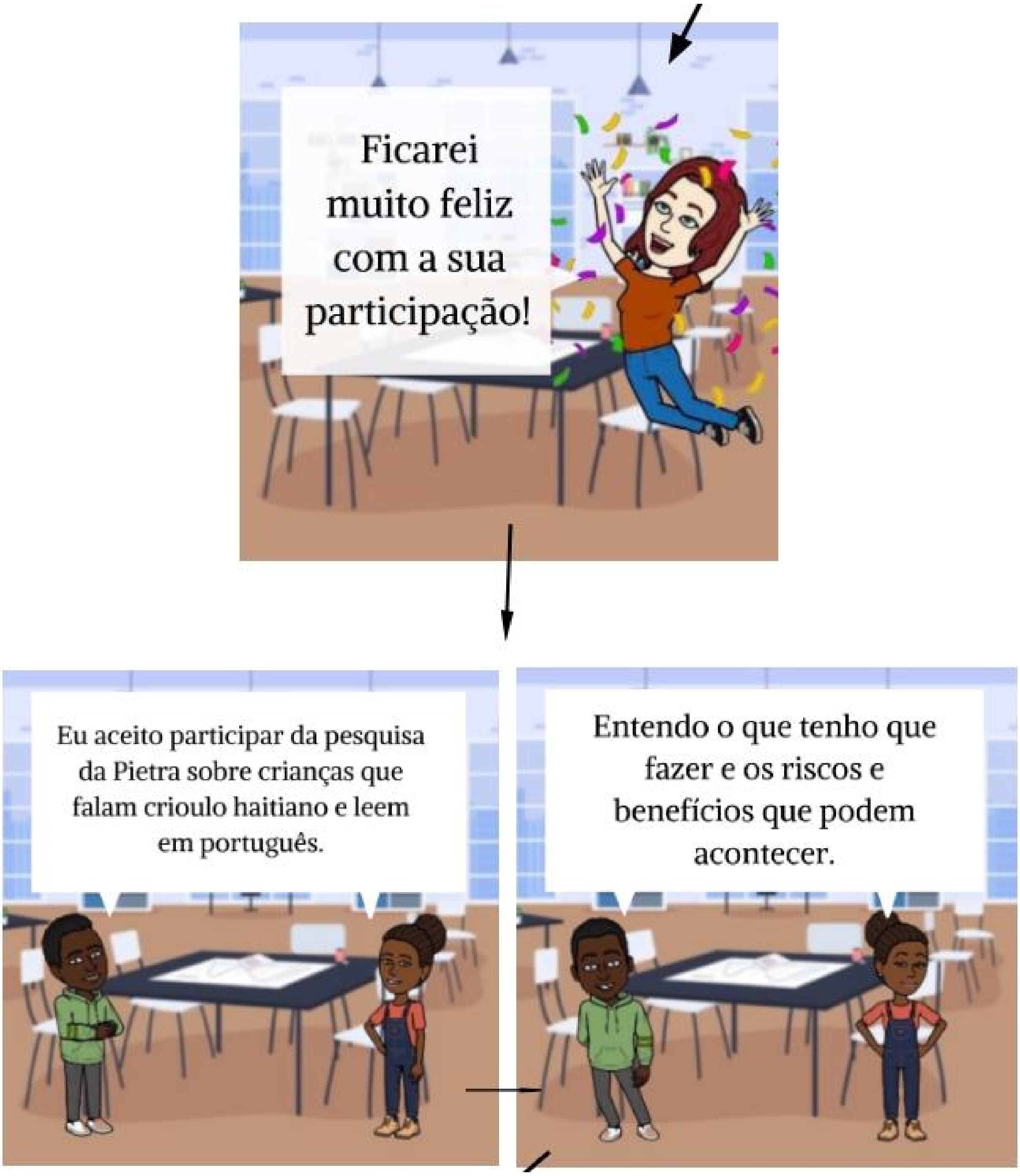
UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS
LABORATÓRIO DA LINGUAGEM E PROCESSOS COGNITIVOS

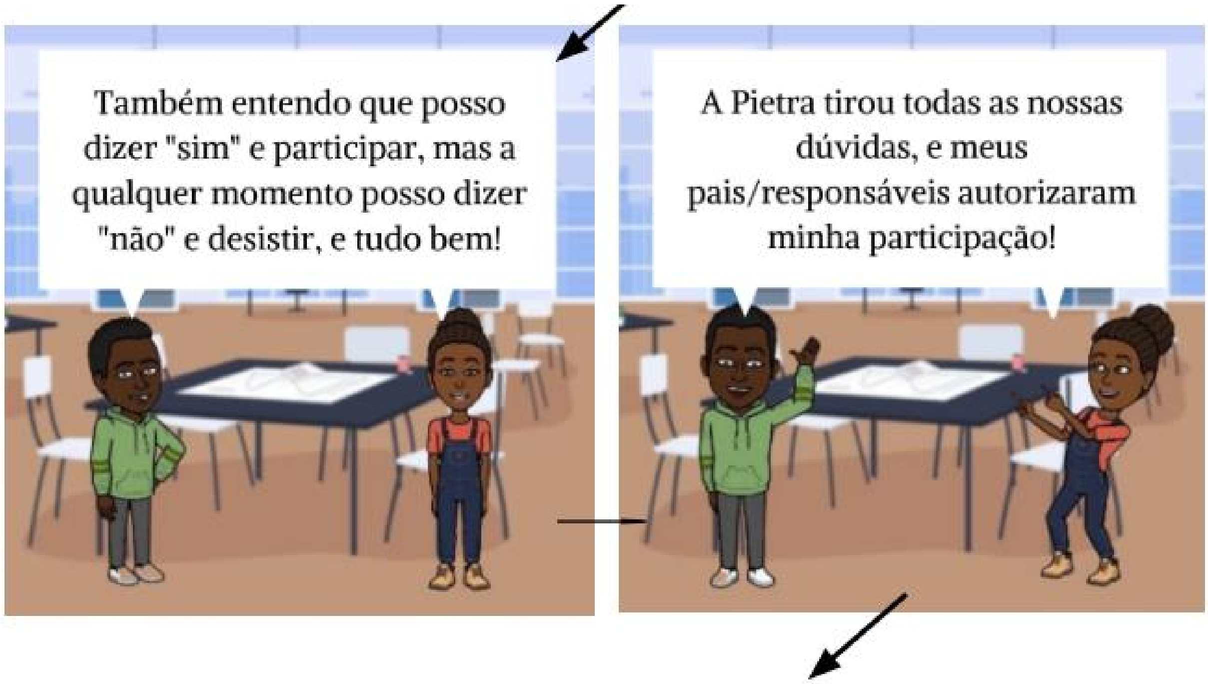
TERMO DE ASSENTIMENTO LIVRE E ESCLARECIDO

baseado na Resolução 510/16 do CNS (Conselho Nacional de Saúde)


Se você aceitar, você vai fazer algumas tarefas no computador.
Você vai ler e ouvir palavras e dizer se elas são palavras mesmo.
Também vamos brincar com o som das palavras e vamos ver letras do alfabeto.
Você vai dizer se conhece algumas palavras em crioulo haitiano.


Seus pais vão responder algumas perguntas sobre quais línguas você fala e quando e com quem fala.
Se você se sentir cansado, chateado ou nervoso, você pode me avisar! Podemos parar a tarefa e continuar depois. E você pode desistir se quiser, tudo bem!
É importante você saber que seu nome, idade ou qualquer outra informação sua desse tipo não será compartilhada.
?
Você tem alguma pergunta? Você pode tirar qualquer dúvida comigo!


Ficarei muito feliz com a sua participação!
Eu aceito participar da pesquisa da Pietra sobre crianças que falam crioulo haitiano e leem em português.
Entendo o que tenho que fazer e os riscos e benefícios que podem acontecer.



( ) Eu aceito participar desta pesquisa.

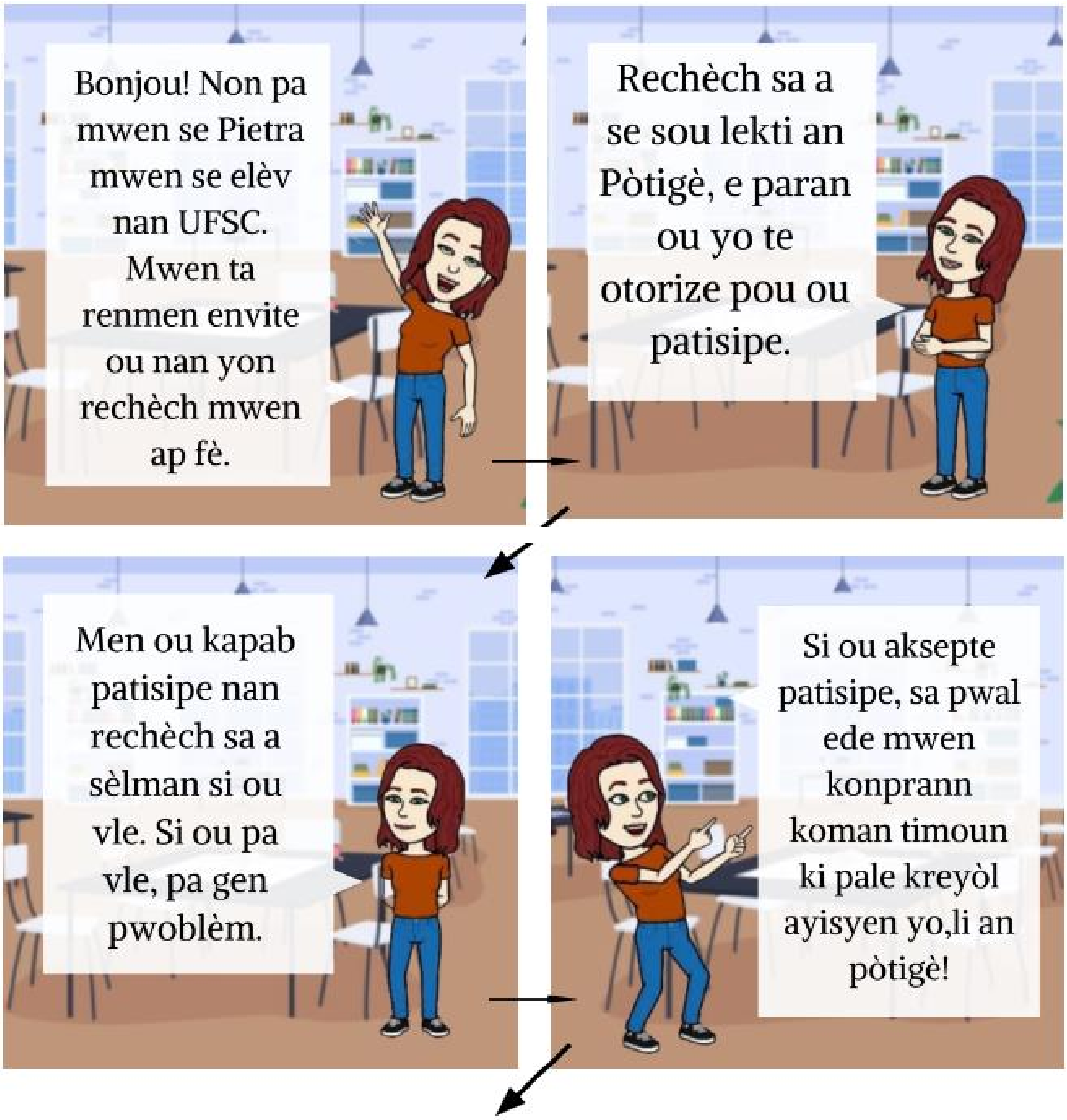
( ) Eu não aceito participar desta pesquisa.

APPENDIX F – Willingness form in HC used in Study 1

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INGLÊS
LABORATÓRIO DA LINGUAGEM E PROCESSOS COGNITIVOS

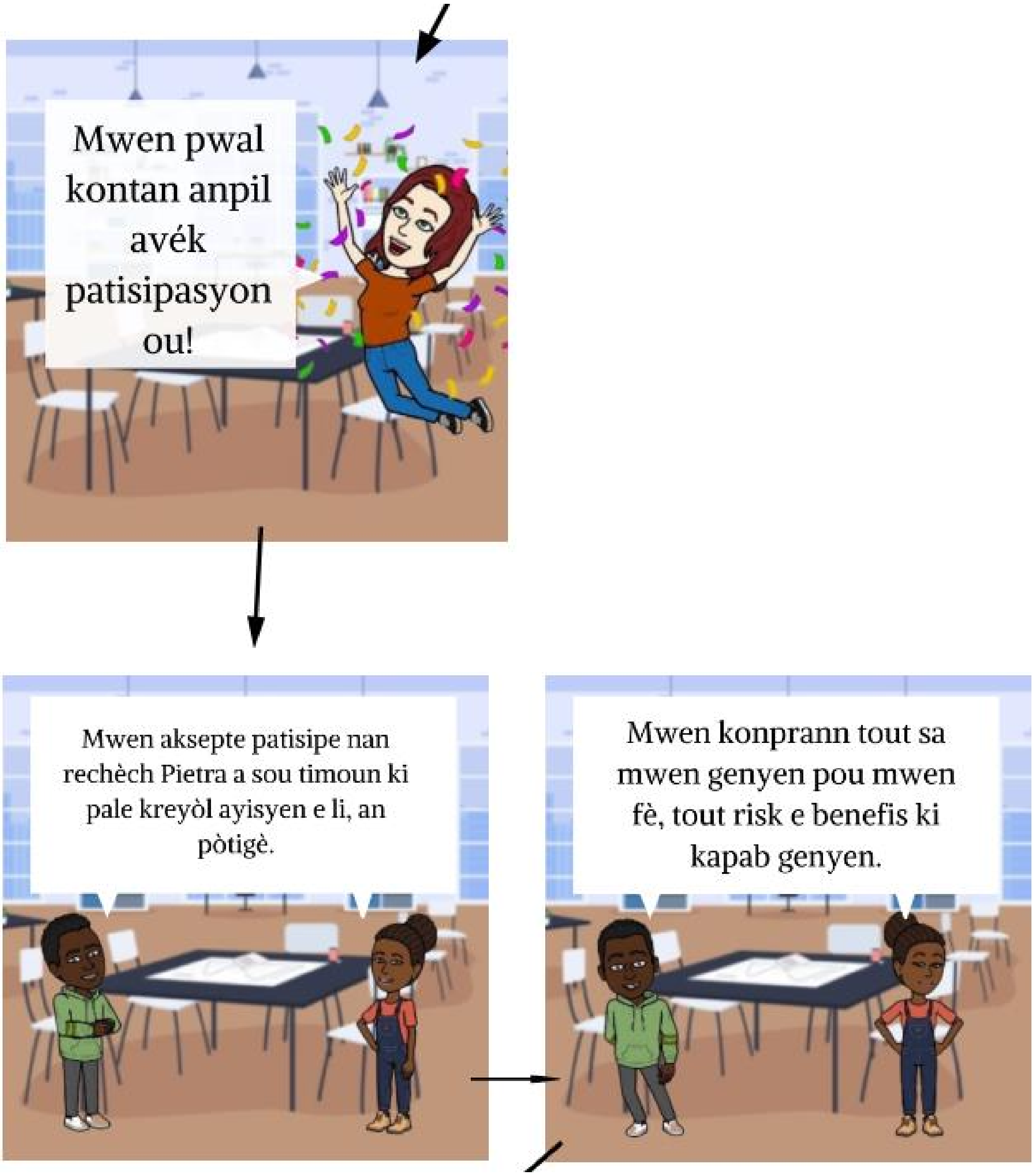
TÈM DE KONSANTMAN LIB E KLÈ

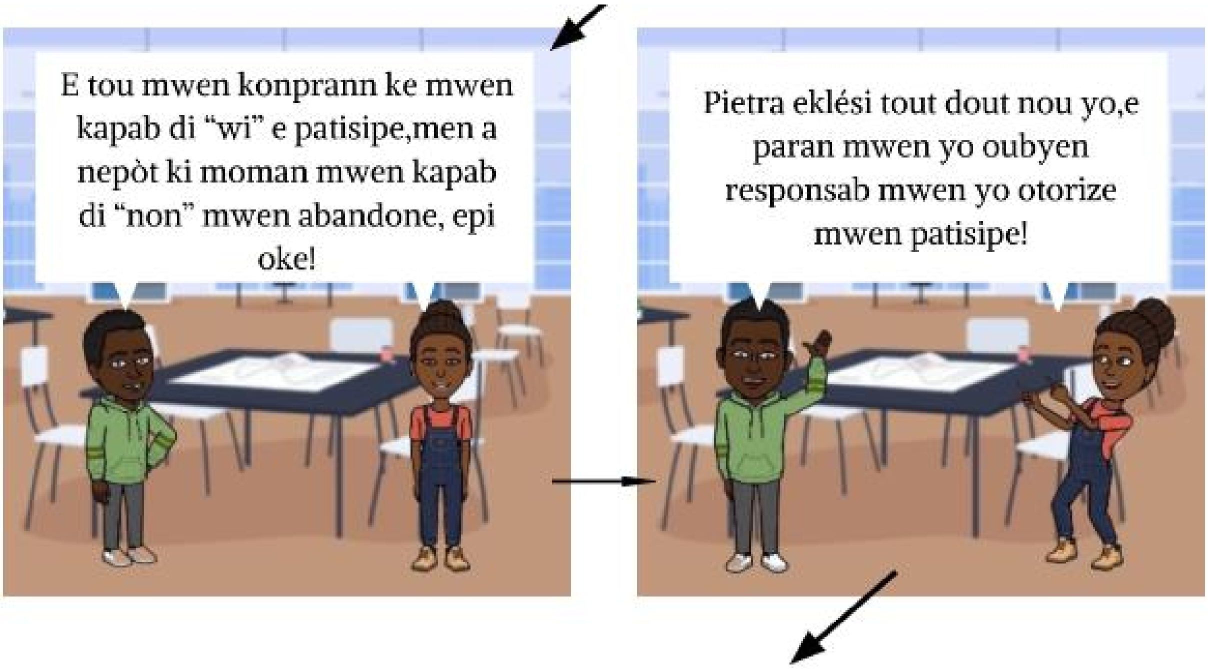
baseado na Resolução 510/16 do CNS (Conselho Nacional de Saúde)




Si ou aksepte,ou pwal fè kék devwa sou òdinatè.
Ou pwal li e tande mo yo e ou pwal di si se mo yo ye vre.
E tou nou pwal jwe avèk son mo sa yo e nou pwal wè lèt yo nan alfabè a.
Ou pwal di non kèk imaj an kreyòl ayisyen.


Paran ou yo pwal reponn kék kesyon sou ki lang ou pale ,kilè e avék ki moun ou pale yo.
Si ou santi ou bouke,ou fache oubyen nève,ou ka avèti mwen sa! Nou kapab kanpe devwa a e kontinye apre. E ou kapab abandone si ou vle, oke pa gen pwoblèm!
Se enpòtan pou ou konnen ke non ou, laj ou, lekòl ou, nepòt lòt enfomasyon ou ankò pa pwal pataje.
Ou genyen kèk kesyon? Ou kapab eklèsi nepót dout ou genyen avèk mwen!


Mwen pwal kontan anpil avék patisipasyon ou!
Mwen aksepte patisipe nan rechèch Pietra a sou timoun ki pale kreyòl ayisyen e li, an pòtigè.
Mwen konprann tout sa mwen genyen pou mwen fè, tout risk e benefis ki kapab genyen.



( ) Mwen akepte patisipe nan rechèch sa a.

( ) Mwen pa aksepte patisipe nan rechèch sa a.

APPENDIX G – Reading habits questionnaire in BP used in Study 2

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA
LABORATÓRIO DA LINGUAGEM E PROCESSOS COGNITIVOS

QUESTIONÁRIO DE HÁBITOS DE LEITURA EM PORTUGUÊS BRASILEIRO

Pesquisa: **INFLUÊNCIAS DA COATIVAÇÃO DA L1 NA LEITURA EM L2 POR BILÍNGUES CRIOULO HAITIANO-PORTUGUÊS BRASILEIRO**

Marque a resposta a resposta que indica aproximadamente a frequência com que você lê os seguintes materiais (impressos ou virtuais).

Atualmente, com que frequência você lê estes materiais impressos em português?

	Nunca	Raramente	Uma vez por semana	Alguns dias da semana	Todos os dias
Revistas					
Jornais					
Livros					
Redes sociais					

Atualmente, com que frequência você lê estes materiais virtuais em português?

	Nunca	Raramente	Uma vez por semana	Alguns dias da semana	Todos os dias
Revistas					
Jornais					
Livros					
Redes sociais					

Há algum outro material não citado aqui que você leia em português? Se sim, com que frequência?

APPENDIX H – Invitation poster used in both studies

APPENDIX I – Consent form in BP used in Study 2

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE COMUNICAÇÃO E EXPRESSÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA
LABORATÓRIO DA LINGUAGEM E PROCESSOS COGNITIVOS

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO – TCLE
baseado na Resolução 510/16 do CNS (Conselho Nacional de Saúde)

Pesquisa: **INFLUÊNCIAS DA COATIVAÇÃO DA L1 NA LEITURA EM L2 POR BILÍNGUES CRIOULO HAITIANO-PORTUGUÊS BRASILEIRO**

Prezado(a) participante,

Eu, Pietra Cassol Rigatti, aluna de Doutorado do Programa de Pós-Graduação em Linguística sob orientação da Prof.ª Dr.ª Mailce Borges Mota, na Universidade Federal de Santa Catarina (UFSC), venho por meio deste convidá-lo(a) a colaborar com minha pesquisa.

O objetivo geral desta pesquisa é investigar a influência do crioulo haitiano como primeira língua (L1) na leitura em português brasileiro como segunda língua (L2) em crianças e adultos falantes nativos de crioulo haitiano.

Por isso, gostaríamos de lhe convidar para participar como voluntário(a) deste estudo. Você será solicitado(a) a realizar as seguintes atividades:

1. **Duas tarefas auto-monitoradas com frases em português, uma com frases escritas e outra com frases faladas:** São tarefas em que você vai ler ou ouvir uma frase, uma palavra de cada vez, e responder a perguntas simples sobre essas frases.

2. **Uma tarefa de leitura de palavras em português:** Nesta tarefa, você vai decidir se uma série de letras é uma palavra real ou se não é uma palavra.

3. **Questionário sobre seus hábitos de leitura em português:** Trata-se de cinco perguntas sobre os tipos de materiais que você lê em português e com que frequência você os lê.

4. **Uma tarefa de vocabulário em crioulo haitiano:** Nesta tarefa, você vai ouvir algumas palavras em crioulo haitiano e vai decidir se as conhece ou não.

5. **Questionário sobre uso de línguas:** Por fim, responderá a um questionário em português sobre quais línguas você utiliza em locais diferentes e com pessoas diferentes.

A sessão de participação e as tarefas serão realizadas de forma remota, por plataformas como o Formulário Google, Google Meet e o Zoom. A realização destas atividades pode lhe causar algum desconforto, tédio, nervosismo ou cansaço físico. Para evitar que as atividades sejam cansativas ou desconfortáveis, garantiremos intervalos entre os testes. A sessão de participação deve demorar em torno de 1h30. Garantimos que você pode desistir a qualquer momento sem prejuízo de qualquer natureza para você.

Sua participação não implicará prejuízos ou divulgação de nomes ou identificação dos participantes de qualquer forma. Cada participante receberá um código de identificação, de forma a deixar a sua identidade anônima na pesquisa. Informações como nome e idade serão registradas apenas com o intuito de controle da pesquisadora. Mesmo que não seja a vontade da pesquisadora, pode acontecer de outras pessoas terem acesso às respostas e informações pessoais dos participantes. Para evitar que isso aconteça, as atividades serão realizadas individualmente com cada participante, e as respostas serão usadas apenas para fins de pesquisa científica e ficarão armazenadas anonimamente em um dispositivo acessível somente por senha pessoal da pesquisadora responsável.

Há riscos leves ao participar desta pesquisa. Esses riscos são relacionados a situações de avaliação, como sensação de nervosismo, aborrecimento e/ou constrangimento. Também pode ocorrer cansaço físico, já que a sessão de coleta de dados pode durar de 45min a 1h30min e ocorrerá majoritariamente em frente a uma tela de computador. Para evitar esses riscos, haverá intervalos programados para descansar e se hidratar. Lembramos também que é possível desistir a qualquer momento da pesquisa, sem qualquer prejuízo. Participar da pesquisa não oferece benefícios diretos aos participantes, apenas contribui para a ciência brasileira.

Os resultados desta pesquisa poderão ser divulgados em eventos ou publicações científicas e disponibilizados anonimamente em plataformas seguras, sem revelar nenhuma informação pessoal que possa lhe identificar. Você pode receber os resultados a qualquer momento, é só entrar em contato com as pesquisadoras. Você apenas poderá receber ressarcimento caso haja eventuais despesas em decorrência da participação na pesquisa. Caso isso ocorra, será devolvido o valor que você gastou. Se você tiver prejuízos por causa da pesquisa, você tem direito à indenização.

Nós, pesquisadoras, nos comprometemos a realizar a pesquisa de acordo com a Resolução do Conselho Nacional de Saúde no 510, de 07 de abril de 2016, que estabelece as normas éticas para as pesquisas em Ciências Humanas e Sociais. O Comitê de Ética em Pesquisas com Seres Humanos da Universidade Federal de Santa Catarina (CEPSH-UFSC) é o responsável por aprovar esta pesquisa. O CEPSH-UFSC foi criado para defender os seus direitos e garantir que eles sejam respeitados e que a pesquisa seja realizada de forma ética, assegurando todos os seus direitos e bem estar.

Informo que você tem a garantia de acesso, em qualquer etapa do estudo, a qualquer esclarecimento sobre o estudo. Se tiver alguma consideração ou dúvida sobre a pesquisa, entre em contato pelo e-mail: pietracr@gmail.com; ou pelo telefone (51)99545-0025; ou também com a Prof.ª Dr.ª Mailce Borges Mota através do e-mail mailce.mota@ufsc.br. Você também pode entrar em contato com o Comitê de Ética em Pesquisa com Seres Humanos da UFSC através do e-mail cep.propesq@contato.ufsc.br, do telefone (48) 3721-6094, da página da internet https://cep.ufsc.br/contato/ ou no seguinte endereço: Rua Desembargador Vitor Lima, n° 222, 4° andar, sala 401, bairro Trindade, CEP 88040-400, Florianópolis, Santa Catarina.

Como informado acima, é garantida a liberdade da retirada de consentimento a qualquer momento, e você pode deixar de participar do estudo, sem qualquer prejuízo ou punição. A retirada pode ser solicitada diretamente com a pesquisadora, por telefone ou e-mail. Este documento será assinado em duas vias, uma das quais permanecerá em sua posse para ser guardada.

Por favor escolha uma das opções:

( ) Concordo voluntariamente participar da pesquisa **INFLUÊNCIAS DA COATIVAÇÃO DA L1 NA LEITURA EM L2 POR BILÍNGUES CRIOULO HAITIANO-PORTUGUÊS BRASILEIRO,** de autoria de Pietra Cassol Rigatti. Declaro que li este documento e que compreendi as informações do Termo de Consentimento Livre e Esclarecido. Eu compreendo meus direitos como participante e consinto em participar deste estudo e em ceder meus dados para a pesquisa. Compreendo o objetivo do estudo, a forma como ele será realizado e a possibilidade de retirar o meu consentimento a qualquer momento, antes ou durante a pesquisa, sem penalidade, prejuízo ou perda de qualquer benefício que eu possa ter adquirido.

( ) Eu não consinto em participar da pesquisa.

APPENDIX J – Consent form in HC used in Study 2

INIVÈSITE FEDERAL SANTA CATARINA
SANT DE KOMINIKASYON E EKSPRESYON
PWOGRAM APWÈ-GRADYASYON AN LENGWISTIK
LABORATWA AN LANG E PWOSESIS COGNITIF (MANTAL)

TÈM CONSANTMAN LIB E EKLÈSISMAN – TCLE
Baze nan Rezolisyon 510/16 do CNS (Konsèy Nasyonal Sante)

Rechèch: **ENFLIYANS NAN KOAKTIVASYON L1 NAN LEKTI AN L2 POU BILENG KREYÒL AYISYEN-PÒTIGÈ BREZILYEN**

Chè Patisipan,

Mwen menm, Pietra Cassol Rigatti, elèv kap fè Metriz nan Pwogram Apwè Gradyasyon an Lengwistik sou oryantasyon Prof.ª Dr.ª Mailce Borges Mota, nan Inivèsite Federal Santa Catarina (UFSC),Mwen vini pou mwen mande ou pèmisyon pou envite pitit ou a kolabore avèk rechèch mwen an.

Objektik general rechèch sa a se mennen ankèt sou enflyans kreyòl ayisyen kòm premye lang (L1) nan lekti an pòtigè brezilyen kòm dezyèm lang (L2) pou timoun e jèn ki pale kreyòl ayisyen natif natal.

Se pou sa,nou ta renmen envite pitit ou a pou li patisipe kòm volontè nan etid sa a,avèk konsantman ou. Nou pwal mande pou Pitit ou a reyalize aktivite sa yo:

1. **De (2) devwa kap sou kontwòl avèk plizyè fraz an pòtigè,yonn avèk fraz ekri e lót yo avèk fraz:** Se devwa kote ou pwal li ou tande yon fraz ,yon mo a chak fwa,epi reponn ak plizyè kesyon sinp sou fraz as yo.
2. **Yon devwa li plizyè mo an:** Nan devwa sa a,ou pwal deside sou yon seri de lèt si se mo yo ye vre oubyen si yo pa mo vre.
3. **Kesyonè sou kisa ou abitye li an pòtigè:** Yo trete de senk kesyon sou ki tip materyél ke ou li an pòtigè e avèk ki frekans ou li yo.
4. **Yon devwa vokabilè an kreyòl ayisyen:** Nan devwa sa a, ou a pwal tande kèk mo an kreyòl Ayisyen epi ou pwal deside si ou konnen yo oubyen non.

5. **Kesyonè sou itilizasyon langaj:** Pou fini,ou pwal reponn yon kesyonè an pòtigè sou ki lang ou itilize nan lokal diferan e avèk moun diferan.

Sesyon patisipasyon an e devwa yo pwal realize sou fòm online,nan platafòm tankou Fòmilè Google e nan Zoom.Reyalizasyon aktivite sa yo kapab fè ou a santi kèk ti malèz,annwi,nève ou bouke fizikman. Pou evite ke aktivite sa yo pa bay fatig oubyen malèz,nou garanti nou ke ap genyen entèval tan rekreyasyon nan tès yo. Sesyon patisipasyon an dwe dire 1h. Nou garanti ou,ke ou kapab abandone a nepòt moman na ou pap pèdi anyen.

Patisipasyon ou pap pwal enplike okenn pèt oubyen pibliye non ou oubyen idantifikasyon patisipan yo, sou nepòt fòm. Chak patisipan pwal resevwa yon kòd idantifikasyon, nan fòm pou nou kite idantite ou anonim nan rechèch sa. Enfòmasyon tankou non ak laj ou a pwal anrejistre jis pou moun kap fè rechéch la fè kontwòl. Menm si se pa volonte moun kap fè rechèch la ,sa ka rive ke lòt moun gen aksè avèk repons enfómasyon pèsonel patisipan yo. Pou evite sa pa rive ,aktivite yo pwal reyalize endividyèlman avèk chak patisipan,epi repons yo pwal itilize sélman nan fen rechèch syantifik e yo pwal mete yo nan yon dispozitif ki gen aksè sélman avèk kòd pèsonèl moun ki responsab fè rechéch la.

Gen ti risk tou piti nan patisipasyon rechèch sa a.Risk sa yo se risk ki rive nan sitiyasyon evalyasyon,tankou sansasyon nève ,kontraryete,oubyen anbarasman.Kapab genyen fatig fizikman ,paske kolesyon done yo kapab dire 45 minit a 1h30 minit e sa pwal fèt majoritèman devan yon ekran òdinatè ni prezansyèl oubyen online. Pou evite risk sa yo,pwal genyen entèval tan pwograme pou pran ti repo ,bwè dlo pou idrate nou . Sonje tou li posib pou ou abandone a nepòt ki moman nan rechèch la,san okenn pèt . Patisipe nan rechèch sa a pa ofri okenn benefis dirèk a patisipan yo,sèlman pou kontribye nan syans brezilyen an .

Rezilta rechèch sa a kapab ale pibliye nan evènman oubyen nan piblikasyon syantifik e disponibilize nan platfòm ki an sekirite,san revele okenn enfòmasyon pèsonèl ki kapab idantifye ou. Ou kapab resevwa rezilta yo a nepòt kèl moman,se sèlaman antre an kontak avèk moun kap fè rechèch yo. Ou menm oubyen pitit ou a kapab resevwa ranbousman an ka si gen yon evantyèl depans a koz patisipasyon ou nan rechèch sa. An ka ke sa rive vre ,n´ap remèt ou oubyen pitit ou a valè ki depanse a . Si fè pèt akoz rechèch sa,nou genyen dwa pou yo dedomaje ou.

Nou menm, moun kap fè rechèch, nou pwomèt reyalize rechèch la an akò avèk Rezolisyon Konsèy nasyonal sante nan 510, 07 avril 2016, ki etabli nòm ak etik pou rechèch an Syans

Imèn e Sosyal. Komite Etik an Rechèch avèk Èt Imen nan Invèsite Federal Santa Catarina (CEPSH-UFSC) se responsab pou apwouve rechèch sa. CEPSH-UFSC te kreye pou defann dwa ou yo e garanti ke yo respekte e tou pou rechèch yo reyalize nan fòm etik,an sekirite, a tout dwa yo e byennèt sosyal.

Mwen ap enfòme ou ke ou genyen garanti ak aksè ,nan nepòt etap etid sa ,nepòt eklèsisman sou etid sa.Si ou genyen kèk konsiderasyon oubyen dout sou rechèch la ,antre an kontak sou e-mail: pietracr@gmail.com; oubyen sou telefòn (51)99545-0025; oubyen tou avèk Prof.ª Dr.ª Mailce Borges Mota atravè e-mail: mailce.mota@ufsc.br. Ou kapab antre na kontak tou avèk Komite Etik an rechèch avèk Èt Imen nan UFSC atravè e-mail cep.propesq@contato.ufsc.br, nan telefòn (48) 3721-6094, nan paj entènèt https://cep.ufsc.br/contato/ oubyen nan adrès sa: Rua Desembargador Vitor Lima, n° 222, 4° etaj, sal 401, bairro Trindade, CEP 88040-400, Florianópolis, Santa Catarina.

Menm jan sa enfòme anlè a,li garanti ou libète pou ou retire konsantman ou a nepòt ki moman e ou a kapab abandone e pa patisipe nan etid la ankò san okenn pèt ou pinisyon. Demand sa a kapab fèt dirèk avèk moun kap fè rechèch la ,nan telefòn oubyen sou e-mail. Dokiman sa a pwal siyen an 2 kopi,yonn ladan yo pwal rete pèmanan nan men responsab la ,ki genyen pou dwa sere li.

Tanpri chwazi yon opsyon:

( ) Mwen dakò volontèman patisipe nan rechèch sa **ENFLIYANS NAN KOATIVASYON L1 NAN LEKTI AN L2 POU BILENG KREYÒL AYISYEN-PÒTIGÈ BREZILYEN,** nan patènite Pietra Cassol Rigatti. Deklare ke mwen li dokiman sa e mwen konprann enfòmasyon yo tèm konsantman lib e eklèsisman. Mwen konprann dwa mwen yo kóm patisipan nan etid rechèch sa e bay done mwen nan rechèch sa.Mwen konprann objektif etid la ,sou fòm li pwal reyalize e posibilite retire konsantman mwen a nepòt ki moman ,avan oubyen pandan rechèch la ,san penalite ,san pèdi okenn benefis ke mwen kapab jwenn.

( ) Mwen pa dakò patisipe nan rechèch la.

APPENDIX K – Data collection notes file for Study 1

UNIVERSIDADE FEDERAL DE SANTA CATARINA

LabLing

LABORATÓRIO DA LINGUAGEM E PROCESSOS COGNITIVOS

CADERNO DE REGISTRO DE COLETA DE DADOS

Projeto: INFLUÊNCIAS DA COATIVAÇÃO DA L1 NO PROCESSAMENTO DE L2 POR BILÍNGUES CRIOULO HAITIANO-PORTUGUÊS BRASILEIRO

ESTUDO 1: EXPERIMENTO ON-LINE EM ESCOLAS

Pietra Cassol Rigatti

Agosto a Dezembro de 2022

Projeto: Influências da coativação da L1 no processamento de L2 por bilíngues crioulo haitiano-português brasileiro. Estudo 1.

Nome do participante	Escola	Turma	Código
			5-
			letter
			and
			5-
			number
			random
			codes

Projeto: Influências da coativação da L1 no processamento de L2 por bilíngues crioulo haitiano-português brasileiro. Estudo 1.

Instruções para o(a) experimentador(a):

Agradecer o interesse do(a) participante. Lembrar que a pessoa pode fazer pausa para descansar e tomar água e que pode desistir a qualquer momento se quiser. Nas tarefas de decisão lexical falada, vocabulário em crioulo haitiano e consciência fonológica, usar **fones de ouvido** se possível. Iniciar pelas **tarefas de decisão lexical** (a não ser que o(a) participante esteja nos anos iniciais, neste caso iniciar pela **tarefa de identificação de letras**). Depois partir para a **tarefa de consciência fonológica** e, por fim, a de **vocabulário em crioulo haitiano**. **Maiores de 11 anos** podem tentar responder o **questionário** ao final se estiverem confortáveis com isso.

Participante nº: __________

Data da coleta de dados: _____/_____/2022

Horário de início do experimento: ____:____ Horário de término: ____:____

Procedimentos realizados pelo participante:

- [ ] Leitura do TALE e aceite de participação
- [ ] Tarefa de decisão lexical falada
- [ ] Tarefa de decisão lexical escrita
- [ ] Tarefa de consciência fonológica
- [ ] Tarefa de vocabulário em crioulo haitiano
- [ ] Tarefa de identificação de letras
- [ ] Questionário de histórico linguístico (responsáveis ou maiores de 11 anos)

Observações: *(registre aqui todo e qualquer evento que ocorrer durante a coleta de dados que achar relevante e que fugir da sequência normal de atividades. Por exemplo: comentários do participante com relação aos estímulos ou à tarefa, problemas com o software/plataforma, etc.).*

____

____

____

____

____

____

____

Projeto: Influências da coativação da L1 no processamento de L2 por bilíngues crioulo haitiano-português brasileiro. Estudo 1.

Instruções para o(a) experimentador(a):

Agradecer o interesse do(a) participante. Lembrar que a pessoa pode fazer pausa para descansar e tomar água e que pode desistir a qualquer momento se quiser. Nas tarefas de decisão lexical falada, vocabulário em crioulo haitiano e consciência fonológica, usar **fones de ouvido** se possível. Iniciar pelas **tarefas de decisão lexical** (a não ser que o(a) participante esteja nos anos iniciais, neste caso iniciar pela **tarefa de identificação de letras**). Depois partir para a **tarefa de consciência fonológica** e, por fim, a de **vocabulário em crioulo haitiano**. **Maiores de 11 anos** podem tentar responder o **questionário** ao final se estiverem confortáveis com isso.

Participante nº: __________

Data da coleta de dados: _____/_____/2022

Horário de início do experimento: ____:____ Horário de término: ____:____

Procedimentos realizados pelo participante:

- [ ] Leitura do TALE e aceite de participação
- [ ] Tarefa de decisão lexical escrita
- [ ] Tarefa de decisão lexical falada
- [ ] Tarefa de vocabulário em crioulo haitiano
- [ ] Tarefa de consciência fonológica
- [ ] Tarefa de identificação de letras
- [ ] Questionário de histórico linguístico (responsáveis ou maiores de 11 anos)

Observações: *(registre aqui todo e qualquer evento que ocorrer durante a coleta de dados que achar relevante e que fugir da sequência normal de atividades. Por exemplo: comentários do participante com relação aos estímulos ou à tarefa, problemas com o software/plataforma, etc.).*

____

____

____

____

____

____

____

APPENDIX L – Data collection notes file for Study 2

UNIVERSIDADE FEDERAL DE SANTA CATARINA

LabLing

LABORATÓRIO DA LINGUAGEM E PROCESSOS COGNITIVOS

CADERNO DE REGISTRO DE COLETA DE DADOS

Projeto: INFLUÊNCIAS DA COATIVAÇÃO DA L1 NO PROCESSAMENTO DE L2 POR BILÍNGUES CRIOULO HAITIANO-PORTUGUÊS BRASILEIRO

ESTUDO 2: EXPERIMENTO ON-LINE COM ADULTOS

Pietra Cassol Rigatti

Agosto a Dezembro de 2022

Projeto: Influências da coativação da L1 no processamento de L2 por bilíngues crioulo haitiano-português brasileiro. Estudo 2.

Nome do participante	Contato	Código
		5-
		letter
		and
		5-
		number
		random
		codes

Projeto: Influências da coativação da L1 no processamento de L2 por bilíngues crioulo haitiano-português brasileiro. Estudo 2.

Instruções para o(a) experimentador(a):

Agradecer o interesse do(a) participante. Lembrar que a pessoa pode fazer pausa para descansar e tomar água e que pode desistir a qualquer momento se quiser. Nas tarefas de compreensão de frases faladas e vocabulário em crioulo haitiano, usar **fones de ouvido** se possível. Iniciar pela **tarefa de decisão lexical escrita**, depois seguir para as **tarefas de compreensão de frases**. Depois, partir para o **questionário de hábitos de leitura** e a **tarefa de vocabulário em crioulo haitiano**. Por fim, responder o **questionário de histórico linguístico**. Se a pessoa não se sentir confortável em responder uma pergunta específica, basta nos avisar.

Participante nº: __________

Data da coleta de dados: _____/_____/2022

Horário de início do experimento: ____:____ Horário de término: ____:____

Procedimentos realizados pelo participante:

- [ ] Leitura do TCLE e aceite de participação
- [ ] Tarefa de decisão lexical escrita
- [ ] Tarefa de compreensão de frases faladas
- [ ] Tarefa de compreensão de frases escritas
- [ ] Questionário de hábitos de leitura
- [ ] Tarefa de vocabulário em crioulo haitiano
- [ ] Questionário de histórico linguístico

Observações: *(registre aqui todo e qualquer evento que ocorrer durante a coleta de dados que achar relevante e que fugir da sequência normal de atividades. Por exemplo: comentários do participante com relação aos estímulos ou à tarefa, problemas com o software/plataforma, etc.).*

____

____

____

____

____

____

____

Projeto: Influências da coativação da L1 no processamento de L2 por bilíngues crioulo haitiano-português brasileiro. Estudo 2.

Instruções para o(a) experimentador(a):

Agradecer o interesse do(a) participante. Lembrar que a pessoa pode fazer pausa para descansar e tomar água e que pode desistir a qualquer momento se quiser. Nas tarefas de compreensão de frases faladas e vocabulário em crioulo haitiano, usar **fones de ouvido** se possível. Iniciar pela **tarefa de decisão lexical escrita**, depois seguir para as **tarefas de compreensão de frases**. Depois, partir para o **questionário de hábitos de leitura** e a **tarefa de vocabulário em crioulo haitiano**. Por fim, responder o **questionário de histórico linguístico**. Se a pessoa não se sentir confortável em responder uma pergunta específica, basta nos avisar.

Participante nº: __________

Data da coleta de dados: _____/_____/2022

Horário de início do experimento: ____:____ Horário de término: ____:____

Procedimentos realizados pelo participante:

- [ ] Leitura do TCLE e aceite de participação
- [ ] Tarefa de compreensão de frases escritas
- [ ] Tarefa de compreensão de frases faladas
- [ ] Tarefa de decisão lexical escrita
- [ ] Tarefa de vocabulário em crioulo haitiano
- [ ] Questionário de hábitos de leitura
- [ ] Questionário de histórico linguístico

Observações: *(registre aqui todo e qualquer evento que ocorrer durante a coleta de dados que achar relevante e que fugir da sequência normal de atividades. Por exemplo: comentários do participante com relação aos estímulos ou à tarefa, problemas com o software/plataforma, etc.).*

____

____

____

____

____

____

____