UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS

Linha de Pesquisa: PSICOLINGUÍSTICA

Cristiane Ely Lemke

EFFECTS OF BILINGUALISM AND BILITERACY ON THOUGHT ORGANIZATION AND SYNTACTIC COMPLEXITY IN BILINGUAL CHILDREN

Porto Alegre

2022

Cristiane Ely Lemke

EFFECTS OF BILINGUALISM AND BILITERACY ON THOUGHT ORGANIZATION AND SYNTACTIC COMPLEXITY IN BILINGUAL CHILDREN

Tese apresentada como requisito parcial para obtenção do título de Doutor em Letras com ênfase em Psicolinguística na Universidade Federal do Rio Grande do Sul.

Orientador(a): Prof. Dra. Ingrid Finger

Porto Alegre

2022

Cristiane Ely Lemke

EFFECTS OF BILINGUALISM AND BILITERACY ON THOUGHT ORGANIZATION AND SYNTACTIC COMPLEXITY IN BILINGUAL CHILDREN

Tese apresentada como requisito parcial para obtenção do título de Doutor em Letras com ênfase em Psicolinguística na Universidade Federal do Rio Grande do Sul.

Aprovado em: 26 de outubro de 2022.

BANCA EXAMINADORA

_____

Ingrid Finger – Universidade Federal do Rio Grande do Sul

_____

Aline Fay de Azevedo - Pontifícia Universidade Católica do Rio Grande do Sul

_____

Janaína Weissheimer – Universidade Federal do Rio Grande do Norte

_____

Natália Bezerra Mota - Universidade Federal do Rio Grande de Janeiro

ACKNOWLEDGMENTS

To Prof. Ingrid Finger, for guiding me and teaching me during the last two and a half years. You have been much more than my favorite teacher: you have been a true friend, who guided me through paths other than this dissertation. You have helped with this and other important, and even more difficult, matters. Thank you for caring so much. Thank you for your enthusiasm, for always pushing me and reminding me of my potential. I do not have words to express my gratitude.

To Prof. Janaina Weissheimer, for teaching me so much about things that I did not think I was able to learn. Thank you for all the support, the time dedicated to this research and for the incredible journey. I am sure it is not over yet and we will be able to take our studies to many other places, hopefully in person.

To Prof. Natália Mota, for gifting us with her knowledge and kindness. We do not always come across professors that are eager to share so much, inspiring us with your knowledge and commitment to science. I am honored to have met you and anxious to finally meet you in person.

To Prof. Aline Fay, for accepting to read this dissertation and helping me improve it. I'm also very thankful for your help when I was writing my systematic review.

To Prof. Erica Rodrigues, for asking questions and helping us think about linguistic aspects concerning *Speechgraphs*.

To Larissa Cury, my research partner, from whom I have learned so much. You were the person who helped me find my clones and so much more. Thank you for being so detail-oriented and helping me improve the analysis for the dissertation. I would not have done it without your support and dedication. We have only talked on the phone and online, but it feels like you have been by my side the whole time.

To Ana Clara Costa, Ana Paula Scholl, Hanna Kahn and Diêgo Leandro for sharing your knowledge and helping me learn about graphs and T-Units.

To Célia Corrêas dos Reis, my big boss, for all the support during the research and for believing in me and in this study. I hope I can share with you a little of what I have learned as you have always been so kind to teach me everything you know.

To IENH, for supporting research and its teachers. I am proud to be part of this institution.

To all participants who took part in this study. You are the true reason this dissertation exists. Thank you so much for your time, patience and enthusiasm in helping me study and discover more about biliteracy.

To my friends and encouraging workmates, who led me to this doctoral study: Aline Jaeger, Daniele Blos Bolzan and Luciana Brentano. You were the ones who convinced me I could do it and also the ones who fled with me whenever I needed some time off. Thank you so much,

Dani, for all the support reading my project and for pushing me into this. Lu, thank you for being with me during the last 4 years. The journey has been a lot better with you by my side.

To all my friends, who encouraged me but also helped me laugh and have fun during the last 4 and a half years. Thank you for always being there for me. I am glad I have you. Camila Lanius and Paola Fay, thank you for helping me with the kids. It meant a lot to me.

To my niece, Marina, for being a wonderful roommate during the time I had to stay in Porto Alegre.

To my mother and sisters. Each one of you lives in me. You make me the happiest little sister and daughter in the world. Thank you for spoiling me and being there for me at all times, the good and the difficult ones. I could not ask for better people to share my life with and for that I am grateful. I love you!

To my dad, *in memoriam*, for providing me with the best education possible throughout my life and for helping me follow my dreams. I am so proud to be your daughter.

To Marcelo, my husband, for facing good and bad moments with me, for holding my hand and keeping me close, but also for letting me fly and follow my dreams. For cheering with my achievements and encouraging me to always do better. Thank you. I love you!

To Lara and Cauê, my beloved children, for teaching me more than anyone could. For teaching me and showing me the greatest love of all. You inspire me to be a better person, a better professional and, above all, a better mom. I am sorry for the days and nights I had to study and could not stay with you, but I hope you understand this is also for you. I hope you have learned that we are always to do our best and to follow our dreams, but keeping the people we love close to us. I love you, all the way to the moon and back.

ABSTRACT

Recent studies have shown positive effects of biliteracy on language and cognitive skills in school-age children. However, there is a scarcity of empirical research that portrays the development of writing in bilingual schooled children, specifically in prestigious bilingual schools. Existing research covers more the characteristics of North American and European countries, with extensive experience in bilingual education, but with a different reality from that reported in this dissertation, which is representative of the Brazilian context, in which children have contact with L2, in general, only in the school environment, in which they learn the language and through the language. In this context, the study that is reported in this dissertation, conducted in the field of Psycholinguistics of Bilingualism, had the general objective of investigating, in a transversal and longitudinal way, the effects of schooling in two languages, Portuguese and English, on the production of students from fourth to sixth grades of elementary school. More specifically, we sought to analyze the effects of school bilingualism and biliteracy on the levels of thought organization (connectedness measures) and syntactic complexity in narratives produced in the children's two languages, Portuguese and English. Two studies were conducted: a pilot study, carried out in 2020, with 50 children in the fifth and sixth grades, focusing on the development of writing in L1 and L2; and a second study, carried out in 2021, based on the results of the pilot study. With 118 children, from the fourth to the sixth year of elementary school, the second research focuses on the development of written production in a transversal and longitudinal way, in addition to presenting data on oral production in both languages. An analysis of the correlation of students' productions in L2 with proficiency was also made, revealing that there is a positive correlation between oral and written production and proficiency in L2. The evaluation of measures of connectivity attributes of thought was done through the analysis of graphs with the computational tool *SpeechGraphs* (Mota et al., 2012, 2016, 2019) and the syntactic complexity was verified through the analysis of T-Units and the Subordination Index (Hunt, 1965). The results indicate that the children's textual productions develop in a crescent along the school years, in a parallel way in the two languages, with an advantage for the texts in Portuguese, L1 of the children. The longitudinal analysis shows the growth of participants from one year to the next in both languages, although a less significant growth is perceived in L2, which could also be interpreted as an impact of the pandemic. Based on the results found, we consider it essential that the development of students' biliteracy be accompanied by teachers in an integral way, taking into consideration the fact that the children are developing academic and linguistic skills in two languages. There is a need for greater investment in research and teacher training to work with bilingual education so that we have greater clarity of the processes involved in the written and oral productions of children in the context of bilingual schooling, in a context in which L2 is not present in the community.

Keywords: Bilingualism. Biliteracy. Thought organization. Syntactic complexity. Graph analysis.

RESUMO

Estudos recentes têm demonstrado efeitos positivos da biliteracia nas habilidades linguísticas e cognitivas das crianças em idade escolar. Verifica-se, contudo, uma escassez de pesquisas empíricas que retratam o desenvolvimento da escrita em crianças bilíngues escolarizadas, especificamente em escolas bilíngues de prestígio. As pesquisas existentes abrangem mais as características de países da América do Norte e da Europa, com ampla experiência em educação bilíngue, mas com uma realidade distinta da reportada nesta tese, que é representativa do contexto brasileiro, em que as crianças têm contato com a L2, de modo geral, apenas no ambiente escolar, no qual aprendem a língua e através da língua. Nesse contexto, o estudo que é relatado nesta tese, conduzido no campo da Psicolinguística do Bilinguismo, teve como objetivo geral investigar, de forma transversal e longitudinal, os efeitos da escolarização em duas línguas, português e inglês, na produção dos alunos de quarto a sexto ano do Ensino Fundamental. Mais especificamente, buscou-se analisar os efeitos do bilinguismo escolar e da biliteracia nos níveis de organização do pensamento (medidas de conectividade) e de complexidade sintática em narrativas produzidas nas duas línguas das crianças, português e inglês. Dois estudos fazem parte desta tese: um estudo piloto, realizado em 2020, com 50 crianças dos quintos e sextos anos, com foco no desenvolvimento da escrita em L1 e L2; e um segundo estudo, feito em 2021, desenvolvido a partir dos resultados do estudo piloto. Contando com 118 crianças, do quarto ao sexto ano do Ensino Fundamental, a pesquisa se volta para o desenvolvimento da produção escrita de forma transversal e longitudinal, além de apresentar dados da produção oral nas duas línguas. Uma análise da correlação das produções dos alunos em L2 com a proficiência também foi feita, revelando que há uma correlação positiva entre produção oral e escrita com proficiência em L2. A avaliação de medidas de atributos de conectividade do pensamento foi feita através da análise de grafos com a ferramenta computacional *SpeechGraphs* (Mota *et al*., 2012, 2016, 2019) e a complexidade sintática foi verificada através da análise de T-Units e do Índice de Subordinação (Hunt, 1965). Os resultados indicam que as produções textuais das crianças se desenvolvem em uma crescente ao longo dos anos escolares, de forma paralela nas duas línguas, com uma vantagem para os textos em Português, L1 das crianças. A análise longitudinal mostra o crescimento dos participantes de um ano para o outro em ambas as línguas, embora perceba-se um crescimento menos significativo na L2, o que pode também ser interpretado como um impacto da pandemia. Com base nos resultados encontrados, consideramos essencial que o desenvolvimento da biliteracia dos alunos seja acompanhado pelos professores de forma integral, considerando-se o fato de que as crianças estão desenvolvendo suas competências linguísticas e acadêmicas em duas línguas. Faz-se necessário um maior investimento em pesquisa e formação de professores para atuar com educação bilíngue a fim de que tenhamos maior clareza dos processos envolvidos na produção escrita e oral de crianças em contexto de escolarização bilíngue, em uma realidade em que a L2 não está presente na comunidade.

Palavras-chave: Bilinguismo. Biliteracia. Organização do pensamento. Complexidade sintática. Análise de grafos.

CIP - Catalogação na Publicação

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AoA	Age of Acquisition
E	Edges
EFs	Executive Functions
L1	First Language
L2	Second Language
LCC	Largest Connected Component
LSC	Largest Strongly Connected Component
N	Nodes
PE	Parallel Edges
RE	Repeated Edges
SI	Subordination Index
SLA	Second Language Acquisition
UFRGS	Universidade Federal do Rio Grande do Sul
UFRN	Universidade Federal do Rio Grande do Norte
WM	Working memory

TABLE OF CONTENTS

1 INTRODUCTION AND STATEMENT OF PURPOSE

Studies on biliteracy have gained increasing attention recently, mainly due to the visibility that child bilingualism and bilingual education has received in the world and also in Brazil in the past decade. There are several communities in the world in which children are growing up bilingual and literate in two or more languages (Reyes, 2012), which has increased the need for more investigation on the topic. Within this context, research in the area has grown considerably due to the need to better understand and improve the learning experiences of school-age children of different linguistic backgrounds. There is real interest in understanding how the development of reading and writing occurs when literacy involves more than one language, as well as in measuring the consequences of such experience on both cognitive and linguistic development in these children.

In our country, the most widespread form of bilingual education is mainly anchored in the principles of prestige or elite bilingualism (Paulston, 1980), which occurs when both languages are valued at school and in the community. This situation differs significantly from the one in countries where one of the languages is the minority language, for example, Spanish in the United States, where there is a need to ensure respect and appreciation of the language children bring from home to school.

Despite the distance from the reality of countries like the United States compared to Brazil, studies developed in other parts of the world have served as a theoretical basis for discussions and practices of bilingualism, biliteracy, and bilingual education in Brazil, given the theoretical and empirical gap that characterizes our educational context. In this sense, one of the justifications of the present dissertation is precisely the lack of studies related to bilingualism and biliteracy conducted from a Psycholinguistics perspective in Brazil. Psycholinguistics of Bilingualism is an area of research that aims to investigate the underlying linguistic and cognitive mechanisms that make individuals able to learn, process and use more than one language (Finger, 2015). Given the reality of bilingual education that we encounter in Brazil, we believe it is of the essence to understand the process of educating children simultaneously in two languages and the cognitive and linguistic consequences of such a choice in order to analyze and think of better pedagogical approaches that take into consideration this reality.

With regard to school bilingualism in Brazil, unofficial data accounts for more than 500 private schools in the country offering a prestige bilingual curriculum or program. This number does not include bilingual schools for deaf children or for indigenous children[1], which are

---

[1] Indigenous Bilingual Education (Lei de Diretrizes e Bases da Educação Nacional, Law 9,394/1996), Bilingual Education for the Deaf (Law 10,436/2002) Law 14.191, 2021.

guaranteed by law but still exist in small numbers, nor the few public initiatives that are being promoted, as is the case of Bombinhas, in Santa Catarina[2], where the first Portuguese-Spanish bilingual public school was created with the aim of preparing children in the region to work in the tourism sector in the future.

In other words, while there has been a significant increase in the number of bilingual schools in the country, in response to market demands, on the other hand, there is a huge lack of knowledge about what bilingual education is and how to best build a bilingual curriculum that promotes linguistic and cognitive benefits to children. And, for that, we need research that investigates language processing mechanisms that underlie the trajectory these children go through. Many schools have even handed over planning and responsibility for structuring the curriculum to publishers or advisory companies, or they have just extended the number of hours of English language teaching, often after school hours, decoupling these class hours from the rest of the regular school curriculum. In this context, doubts about the best type of teaching methodology to be used in bilingual schools have become increasingly frequent.

For these reasons, research that contributes to a better understanding of the cognitive and linguistic processes that underlie biliteracy from a Psycholinguistics of Bilingualism perspective, considering the typical reality of Brazilian schools, such as the one being proposed here, become even more relevant.

In the case of bilingual children learning to read and write in two languages, we still know very little about how one language system affects the other (Brentano & Finger, 2020), even though it is known that they interact and that there is transfer of knowledge between both the children's languages (Alves & Finger, 2023). We also know that the ability to plan and tell a complex and well-connected story evolves according to schooling and literacy levels (Mota et al., 2016). This is the context of the studies reported in this dissertation, which aimed at investigating the effects of bilingualism and biliteracy on the levels of thought organization (connectedness measures) and on syntactic complexity in written and oral productions in two languages (Portuguese and English) in elementary school children.

More specifically, this dissertation, conducted within the field of Psycholinguistics of Bilingualism, investigated the effects of schooling in two languages, English and Portuguese, on the written and oral productions of students from the fourth to the sixth year of elementary school, cross-sectionally and longitudinally. To that end, we analyzed the effects of school bilingualism and biliteracy on the levels of thought organization (connectedness measures) and syntactic

---

[2] Parecer 2016/200/CEE/SC. Resolução para a oferta de Escola Bilíngue e Internacional em Instituições Escolares de Educação Básica para o Sistema de Ensino do Estado de Santa Catarina.

complexity in students' written narratives, both in Portuguese and in English. The evaluation of measures of connectedness attributes was done through the analysis of graphs, with the computational tool SpeechGraphs (Mota et al., 2014, 2016, 2019), and the syntactic complexity's analysis was based on T-Units (Hunt, 1965) and Subordination Index (SI).

To accomplish such a goal, we conducted two empirical studies involving writing and oral development in two languages and a systematic review on the topic. The systematic review came to meet a literature demand, since few studies in the area were known to us. It became important to carry out a very careful and rigorous literature search, to make sure it would be comprehensive.

The first empirical study, conducted in 2020, during the pandemic in which students were having online classes, focused on the analyses of the writing development in L1 Portuguese and L2 English of 50 children in 5th and 6th grades. In this first study, we decided not to include oral tasks, as it was during the pandemic and, at that moment, online classes were shorter than regular classes, which made the individual collection of oral productions unfeasible. The pilot study was fundamental for the organization of the research study as a whole. Based on the observations during this process, some adjustments were made for the data collection in 2021, which involved a comparison between oral and written production in both L1 Portuguese and L2 English. Also, the possibility of analyzing the data longitudinally represents a great contribution to the area.

The second empirical study was done in order to improve methodological issues of the first, and involved oral production as well as written production. In addition, it covered a greater number of participants (n=118), with children from the fourth to the sixth grade of elementary school. In this study, face-to-face classes had already been resumed, therefore it was also possible to include a cognitive measure as well as a L2 proficiency measure in the protocol.

The systematic review and both empirical studies will be presented in separate chapters of this dissertation. This dissertation is, therefore, divided into 6 chapters, Chapter 1 being this introductory piece. Chapter 2 offers the literature review that provides the basis for the studies that were conducted and that are reported here. The chapter begins with a discussion of the terms 'Bilingualism' and Biliteracy following a characterization of Bilingual Education in the world and specifically in Brazil. It leads to a reflection on the way children learn how to write in two languages, going past different ideas and perceptions on the topic over the years. Following, there is a brief account on the origins of graph theory, a description of its elements and their representations. Chapter 2 ends with a discussion of syntactic complexity focused on the measure of T-Units and Subordination Index (SI), used in the analyses conducted in the studies reported here.

Chapters 3, 4 and 5 are presented as research articles, with 2 of them having already been submitted and accepted for publication. These chapters are organized in a way to show

the trajectory of the present investigation and the steps taken in the verification of the specific objectives presented above. A more detailed overview of each is presented in the next paragraphs.

Chapter 3[3] presents a systematic review of studies on biliteracy published between the years of 2012 and 2021 that have been conducted with school children. The review was based on the search for articles in English, Spanish and Portuguese in two large databases: Institute of Education Sciences (ERIC) and Scientific Electronic Library Online (SCIELO). The objective of this review was to answer two questions: "What studies are being carried out on the development of writing in bilingual children undergoing biliteracy?" and "What are the conclusions of these studies regarding the development of writing skills in two languages?" These answers will be presented in the third chapter of this dissertation and provide the theoretical basis for the investigation of our objectives, considering the lack of studies involving bilingual school-aged children, especially in realities where the L2 is not present in the students' daily lives.

Taking into consideration this empirical gap, chapter 4 presents the Pilot Study held in 2020[4], and which served as the basis for the study that is presented in Chapter 5. The main goal of this study was to investigate the effects of bilingualism and biliteracy on the levels of thought organization (connectedness measures) and syntactic complexity in the written production in Portuguese and English in a group of fifty middle-aged students (Mean age=10.7), enrolled in a bilingual school in the south of Brazil. As the Pilot Study was held during the pandemic, students were having online classes. Analyzing the process undertaken during the Pilot Study was extremely relevant in order to continue our research project the following year.

Chapter 5 reports the study that was conducted in 2021, shedding light to the general discussion proposed in this dissertation. After the pilot study, we made a few adjustments and added materials (oral production, proficiency and cognitive measures were included) to the design of the experiment, and Chapter 5 presents the results of the investigation of the effects of bilingualism and biliteracy on thought organization and syntactic complexity in written and oral data in both English and Portuguese in school-aged children enrolled in a bilingual prestige curriculum. Since we are also interested in the trajectory of language development in such a bilingual education context, this study also focused on the longitudinal analysis of written productions along a two-year period.

Finally, Chapter 6 brings some final remarks, discussing overall results, listing

---

[3] This part of the dissertation is written in Portuguese as it is a study that has been accepted for publication in a Brazilian book.
[4] https://eurokd.com/LTRQ/doi/10.32038/ltrq.2021.26.01.pdf

limitations of the studies conducted and pointing out some pedagogical implications of the results reported in the present dissertation.

2 LITERATURE REVIEW

In this section, the theoretical assumptions underlying the studies reported in the present dissertation are discussed. Initially, the topics of bilingualism and bilingual education are introduced, followed by a brief analysis of bilingual education in Brazil. After that, the chapter presents a discussion related to the development of writing in two languages, since the research goal of the present dissertation is to investigate the effects of schooling in two languages, English and Portuguese, on the written and oral productions of students from the fourth to the sixth year of elementary school.

In the following sections, Graph Theory and the concept of T-Units and Subordination Index are presented, as they provide the basis for the analysis of the data collected in the two empirical studies conducted in this dissertation project. The chapter was built as a way to lay out the theoretical basis for the methodological decisions that were made in the empirical investigations presented in Chapters 4 and 5. Therefore, the way theoretical and methodological issues are intertwined will be unraveled in the sections below and in the next chapters as well.

2.1 BILINGUALISM, BILITERACY, AND BILINGUAL EDUCATION

For years the definitions for the term 'bilingualism' have been discussed and refined. Numerous attempts to characterize what bilingualism is have been made throughout the decades but in fact there is no single definition that fits all realities in which bilingualism takes place. In the section below, a reflection on the topics of bilingualism and biliteracy will be presented. After that, we will approach the concept of bilingual education, discussing the different models that have been already proposed in the literature. Finally, based on the two previous sections, bilingual education in Brazil will be briefly presented with an emphasis on contexts of Prestige bilingual education, which is the circumstance in which the studies reported in the present dissertation have been conducted.

2.1.1 Bilingualism and Biliteracy

One of the first definitions of bilingualism dates from 1933, when Bloomfield stated that bilingualism is the native control of two languages: "...in cases where perfect foreign language learning is not accompanied by loss of mother tongue, the result is the bilingualism, the native control of two languages" (Bloomfield, 1933, p.55-56). This classic definition marked, for a long time, what was expected of a bilingual individual: native control of both languages, that is, a perfect ability, or equivalent/balanced ability, in two, or more, languages of the individual.

However, over time, the definitions went from one extreme to the other. Macnamara opposed this view and proposed that "a bilingual individual is someone who possesses minimal competence in one of the four language skills (speaking, listening, reading and writing) in a language other than their native language" (Macnamara, 1967, in Hamers and Blanc, 2000, p. 7). Such a view expanded and broadened the concept of bilingualism, including as bilingual individuals a much wider range of people, since it considered a minimum competence in any of the skills as a sufficient criterion to position an individual within the bilingual category. Following the same lines, Diebold's (1964) concept of incipient bilingualism categorized individuals with minimal competence - those able to say a few sentences in a language other than their mother tongue - as bilinguals, also moving from an extreme that excluded many speakers, to an extreme that included too many speakers.

Vaid's definition (2002, in Zimmer, Finger, Scherer, 2008, p. 05) can be added to the list above. In addition to level of proficiency, the author emphasizes the use of languages as one of the characteristics of bilinguals, when he defines bilinguals as "individuals who know and use two languages, which would not necessarily be used in the same context, nor mastered at the same level of proficiency". This definition brings *use* as an essential part of understanding what bilingualism really is, since the vision of perfect or balanced bilinguals was, eventually, questioned. Considering that an individual uses his languages in different situations and for different purposes, it was expected that they would hardly have similar levels of skills in both languages, since the use of each can be, and commonly is, different.

In seeking to define bilingualism, Grosjean (1999) draws attention to the intermediate stages between being monolingual and being bilingual along a continuum, highlighting the dynamism of bilingualism, since the profile of the bilingual changes and undergoes alterations over time. The notion of the continuum allows us to see the different degrees of competence of speakers and how much evolution or loss there has been along that continuum.

Along the same lines, Valdés (2003) highlights the continuum between being bilingual and being monolingual. The continuum would be similar to the figure below, with the first letter representing the dominant language and the font size marking the different levels of proficiency:

Figure 1 - A 'bilingual' individual's proficiency
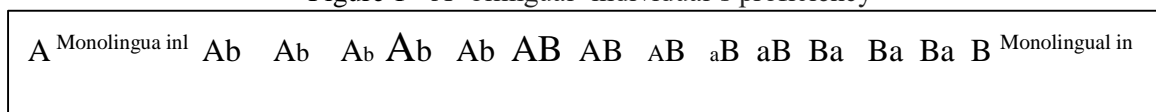
Source: Baker (2006, p. 8)

In this continuum, it is possible to contemplate from speakers at beginner levels to extremely competent and fluent second language users. The different scales within the continuum,

which are not fixed and can undergo modifications along the way, reveal that language knowledge is not static and the mastery and use of them may vary depending on the extent these languages are used everyday throughout a person's life.

Based on this discussion, it is possible to observe that there has been a radical shift in the literature regarding the characterization and definition of bilingualism and the extent that distinct experiences in a person's life may affect his/ her language development. In addition, as a result of experience, levels of proficiency and fluency in the two languages may change throughout time and analyzing these changes may provide a unique lens to understand the mechanisms that underlie language processing and thought (Kroll & Navarro-Torres, 2018).

Regarding the difficulty of selecting a definition of bilingualism that encompasses all the characteristics and possibilities of bilingual individuals, because these are not fixed and immutable, Luk and Bialystok (2013) also point out that there is no clear line between what constitutes a bilingual experience and a monolingual experience. Grosjean (2013) adds that bilingualism is a dynamic and interactive experience, characterized by individual and context factors, corroborating the idea that there is no right or defined point that we can mark as a transition from mono to bilingualism. Surrain and Luk (2017) reflect on the term arguing that not only is the bilingual experience related to the degree of bilingualism, but also to other multiple factors, such as the history and context of acquisition, proficiency and use of each language.

Considering the many variables that influence its characterization, Baker (2006) lists eight dimensions of bilingualism, namely ability, use, language balance, age, development, culture, context and elective bilingualism, which may interact and overlap. The first one, ability, refers to productive competence, that is, the ability to express oneself (speaking and writing) in two languages, as well as to receptive skills, related to being able to understand or read in two languages. These skills are also best placed along a continuum, with the possibility of being more developed in one language more than in the other, as well as varying throughout life, for they are not static.

The second domain, use, is related to the specific use the individuals make of each language, as languages are acquired for a specific use or purpose in a particular context: the language of the family, the school, the television, the street, etc. Some people play, others work, others shop in one language, but read, sing and study in another, for example. Often, each language of an individual is used for different purposes, with distinct people, and/or in different contexts.

As for the balance of languages, Baker claims that rarely the two (or more languages) are fully balanced, being one of them usually dominant. The idea of "perfect" bilinguals, with balanced ability in both languages, and the hypothesis of the "Double Monolingual", also rejected by Grosjean (1997) and Cook (2003), are demystified, as it is quite difficult to achieve full

proficiency in a given language, or in two or more languages, in the four skills (speaking, writing, listening and reading), considering that the use (one of the dimensions mentioned above), the context and the interlocutors vary, as also highlighted by Zimmer, Finger and Scherer (2008).

One of the most mentioned aspects in the literature that intends to provide a definition of bilingualism is related to the age or moment of language acquisition or first exposure. For instance, Baker (2006) considers the process of acquiring two languages from birth as simultaneous or early bilingualism and when the process takes place later (after the age of three, according to him), it is called consecutive or sequential bilingualism. On the other hand, Hamers and Blanc (2000) separate these moments into three: child bilingualism, before 10 or 11 years of age, adolescent bilingualism, between 11 and 17 years of age, and adult bilingualism, after 17 years of age. Hamers and Blanc further subdivide child bilingualism in simultaneous and consecutive ways. Although there are differences among scholars as to the exact age for each case, there is, in all of them, a distinction between simultaneous/early and sequential/consecutive bilingualism.

Regarding the context in which bilingual individuals live, it is observed that some are contexts where the two languages are present in the community on a daily basis, which are normally called endogenous communities. These would be, for example, border regions, multilingual countries, places where immigration languages are present, among others. The most monolingual contexts, where there are few opportunities for a person to interact in a second language in the community are known as exogenous contexts.

Another distinction that is sometimes made is between additive and subtractive bilingualism. Additive bilingualism is the case in which a prestige language is added, with the assumption that this knowledge will come without causing loss or damage to the first language, as occurs in schools of bilingual education of choice. In contrast, subtractive contexts relate to situations in which the country's policy favors the "replacement" of the speakers' first language (usually the family language) by the language of the country in which children are living, the majority language, as is the case of Mexican immigrants in the United States, for example.

Another dimension presented by Baker (2006) is quite relevant for the work with bilingual education, since it deals precisely with elective bilingualism, when there is a choice for learning more than one language. This is the case of students in bilingual schools, for example, who add a language to their repertoire. Also included in this type of bilingual are English-speaking students who study in immersion schools, as in the Canadian context, becoming English/French bilinguals. May (2017) adds that this type of bilingualism is seen as beneficial, both socially and cognitively, as well as educationally. Baker (2006) also describes circumstantial bilinguals, a situation in which they need to learn a new language in order to interact in the society in which they are immersed. Examples of this type of bilingualism are immigrants, who, in order to be able to live and operate

in the country to which they migrated, need to become bilingual. However, in these cases, the second language sometimes endangers the first, becoming a context of subtractive bilingualism. This is the case of immigrants who came to Brazil over a century ago and who, gradually, left their native languages aside to speak Portuguese and live in the society in which they were immersed. In subtractive bilingualism, the individual's mother tongue is often seen as problematic, as in the case of Latino students in the United States, whereas, on the other hand, elective bilingualism is seen as an advantage.

The dimensions presented above exemplify, in a certain way, the difficulty of dealing with the issue of bilingualism, since there is no simple definition for the phenomenon. Each individual has a unique language experience and, at the same time, bilingualism is a plural phenomenon, considering that bilinguals make up practically half of the world's population. There are countless situations in which the use of two or more languages (multilingualism) permeates family, work, school relationships, etc. In short, languages transit in people's routine in a common and natural way, making this a much debated and studied subject, which makes it so present, dynamic and current.

Over the years, there has been a shift in the understanding of bilingualism and more up-to-date research has brought different perspectives on the matter. Even though bilingualism was first seen as a liability (Jones & Stweart, 1951), recent studies suggest that "bilingualism changes language, cognition, and the brain in ways that often benefit bilinguals" (Kroll; Navarro-Torres, 2018, p.1). Most of the studies being conducted in the last two decades have been supportive of bilingualism. Due to the contribution of neuroscience methods, problematic issues, such as adult L2 learning and age of acquisition (AoA), have been revisited. As a matter of fact, "imaging studies have shown that even the briefest exposure to a new language can begin to affect sensitivity to L2 structure and that new L2 phonetic learning may reflect individual differences rather than AoA per se" (Kroll & Navarro-Torres, 2018, p. 2). What is even more surprising is that these studies also show that learning a L2 affects knowledge and representation of the L1, even in early stages of L2 learning and even when learners have not yet achieved high levels of proficiency (Bice & Kroll, 2015; Chang, 2012, 2013; Azevedo et al., 2017). Once more, Grosjean's view that bilinguals are not two monolinguals in one is proved right.

Kroll et al. (2018, p.60) argue that "bilingualism is a complex life experience", and we believe it is indeed. Complex not meaning difficult, but with many different possibilities, which, again, make it a hard task to define the term. Nonetheless, there are some things that we must keep in mind when defining bilinguals and bilingualism. One of them is that "at any moment of language use, a bilingual's languages are active in parallel" (Zirnstein, Bice and Kroll, 2019, p. 35.) What changes is their level of activation, which is related to multiple factors, such as the task,

the levels of proficiency in each language and the environment and contexts of use. Bilinguals who have the same age of acquisition and similar proficiency, for instance, may use these languages differently in their everyday lives.

Kroll and Navarro-Torres (2018) add some important factors to contemplate when defining bilinguals. The authors posit that it is not enough to think about bilinguals as 'young' or 'old', because they "differ in the nature of the two languages spoken, the age at which the L2 is acquired, the level of proficiency in the L2, whether the majority language in the environment is the L1 or the L2, and whether the context itself is bilingual" (p. 3). Taking into account only one of these characteristics ends up not representing bilinguals at its best, and this is one of the reasons why sometimes researchers do not find plausible data when comparing bilinguals to monolinguals, since there are many factors that need to be controlled and analyzed in order to better characterize a group.

Following along the same lines, Bialystok (2021) argues that bilingualism can not be handled as a binary category and, much less, opposed to another oversimplification: monolingualism. The author compares it to a package of Swiss Cheese, "with different manifestations of bilingual experience placing the holes in different places that together define the experience"(p.2). In a package, each slice has holes in different places. Together they end up covering the holes of other slices. As a slice of cheese, each experience of bilingualism is different. Nonetheless, many and varied studies with bilinguals of distinct backgrounds (Luk et al., 2011; Vega-Mendonza et al., 2015; Bialystok et al., 2005; Dash et al., 2019; Verrey et al., 2016) have found a connection between specific aspects of bilingual experience and cognitive outcomes. In research, it is therefore essential that descriptions of bilingual participants in the studies are made in detail so as to bring the specificities of the particular group being studied.

In fact, there seems to be a significant paradigm change in more recent studies in bilingualism, which have started to take more into consideration the participants' individual differences. The long-used opposition between monolinguals and bilinguals is giving space to a more fluid definition, which sees bilingualism as a continuum, making it harder to simply compare bilinguals to monolinguals. Even monolinguals are exposed to other languages and this can also somehow impact their "category", which, again, is not stable. Weissheimer et al. (2021) argue that there has been an increase in the number of studies that are considering the role of individuals' differences in the so-called 'bilingual experience'.

Other authors corroborate with the idea of considering the complexity of the bilingual experiences, taking into account not only how diverse experiences can be but also the cognitive and linguistic demands bilingual individuals face. It is common knowledge that bilinguals deal with cross-language competition as they constantly have their languages activated. However, what

needs to be clarified is that their processing demands will vary based on their particular stage of bilingualism (Salig et al., 2021). Important aspects to consider are, for instance, if they are still acquiring their L2, the languages they already know, their daily language use environment, and their processing demands, which may vary depending on who they are talking to - for example, if it is a monolingual person or someone with whom they can use both languages. These different traits may result in different neural and cognitive outcomes. Based on these different patterns across bilinguals, Salig et al. (2021) consider it inappropriate to treat 'bilingual' as a single, indivisible category. Finally, as we can see, the monolingual/bilingual dichotomy proves not to be sufficient when discussing this topic. A thorough analysis of different language experience needs to be done to better characterize and situate individuals in the continuum of bilingualism.

As we have seen, the term 'bilingualism' has received many distinct definitions over the years. In the same way, 'literacy' and 'biliteracy' have been characterized and understood in diverse forms. In this dissertation, literacy is used to refer to the cognitive process of reading and writing (Alves & Finger, 2023). It is in this perspective that we see biliteracy, as the development of literacy, but in two languages.

Alves and Finger (2023) argue that "in alphabetic writing systems, the child establishes a relationship between the graphic symbols and the distinctive sound elements of their language, in order to invest in the discovery of the alphabetic/representational principles of the system in question". In this vein, biliteracy accounts for the capacity to assimilate and compare similarities and differences between the two languages and make use of transfer processes that underlie the development of reading and writing skills in two languages. Becoming biliterate is part of the process of becoming bilingual, but in a different media. In a nutshell, biliteracy is related to the process of learning to read and write and developing a linguistic capacity to express oneself with written words. As Gort (2019, p. 231) points out, "bilingualism and biliteracy are interrelated language-based processes that develop in a parallel fashion and enrich one another". Due to the transfer processes that naturally occur between the two languages of the bilingual individual, reading and writing in one language enriches reading and writing in another one (Alves & Finger, 2022).

Like bilingualism, biliteracy is also an ongoing process. There are some key processes that are the basis to learning to read and write, such as encoding and decoding. Pérez and Torres-Guzmán (1996, p. 54) define biliteracy as "the acquisition and learning of the decoding and encoding of and around print using two linguistic and cultural systems in order to convey messages in a variety of contexts." Encoding and decoding represent one stage of this process, which is also seen as a continuum by Hornberger (2003), who understands biliteracy as a continuum in which an individual's oral and written language abilities change dynamically throughout a person's life,

depending on a series of elements. Gort (2019, p. 233) refers to this "ongoing, dynamic development of concepts and expertise for thinking, listening, speaking, reading, and writing in two languages" as emergent biliteracy.

Within this dynamic process, the transfer of writing strategies across languages is common and routine, as children experiment with patterns and forms of oral and written language, trying to make sense of the languages they are exposed to. It is the schools' mission to provide students with opportunities to develop their bilingualism and biliteracy at best, promoting practices that are in line with studies in the area. Our next section presents a definition of Bilingual Education and introduces the different models that have been already proposed in the literature.

2.1.2 Bilingual Education

One of the classic definitions, by Anderson and Boyer, presents bilingual education as "instruction in two languages and the use of these two languages as a means of instruction for some or all of the school curriculum" (1970, p. 12). Baker and Prys-Jones (1998, p. 466) also speak of instruction as a hallmark of bilingual education: "[...] bilingual education begins when more than one language is used to teach content (such as Science, Mathematics, Social Studies) rather than just being taught as a subject in its own right". Wright and Baker (2017) align with these definitions adding an important detail: "bilingual education often refers to education in which two or more languages are used to teach and learn in part, most or the entire curriculum" (2017, p. 66). In general, using languages as a means of instruction seems to be a common point between the different definitions of bilingual education.

The expression 'Bilingual Education' has already been used in numerous contexts to refer to distinct educational models, a fact that has brought ambiguity to the term. Wright and Baker (2017) recall that the nomenclature has already been used to name contexts in which students are bilingual, but the school emphasizes only the majority language, so the education itself is not bilingual, but the students are. However, nowadays, the term has been understood and used most of the time to characterize educational contexts where two languages are used for daily instruction in the classroom, in which students become bilingual and biliterate. In this type of educational model, both language learning and content learning through languages are the focus of instruction. Nonetheless, even with this current consensus, there are still several models of bilingual education, with different objectives and purposes. In the next paragraphs, some of them will be reviewed.

Fergunson et al. (1997) list ten objectives of bilingual education, which are sometimes conflicting: (1) assimilate individuals and groups to conventional (majority) society; (2) bring unity to multi-ethnic and multilingual countries; (3) enable people to communicate outside their

country; (4) increase language skills for marketing purposes, such as getting a job; (5) preserve ethnic or religious identity; (6) harmonize different linguistic and political communities; (7) spread the use of the colonization language; (8) strengthen elite groups and preserve their privileged position in society; (9) give equal rights and status to unequal languages; and (10) deepen understanding of language and culture. Wright and Baker (2017) modernize the list by adding two items: (11) preserve a minority language at risk; and (12) increase curriculum achievement and school performance. As the authors themselves point out, these two additions make us realize that bilingual education is also related to society's goals, not just school purposes.

Aside from the possible objectives mentioned above, we can divide bilingual education into two major groups of interest: bilingual education for minority groups - also called 'folk bilingualism' (Fishman, 1976, 1977, 2011), and bilingual education for students from the dominant classes - known as 'elite or prestige bilingualism' (Paulson, 1980). Both models will be better explained below.

Bilingual education for minority groups can be divided into two kinds: maintenance and subtractive. Maintenance bilingualism has the goal of preserving the minority language and adding the majority language to the individual's repertoire, whereas the subtractive type has the aim of ensuring that children from minority groups assimilate the dominant (majority) language, without much concern for maintaining the minority language, thus aiming at monolingualism, not bilingualism (Baker, 2006; Genesee & Fortune, 2014; Wright & Baker, 2017). On the other hand, subtractive bilingualism is imposed by society and the educational system. In these situations, the minority language is often seen as inferior, causing the speakers to feel ashamed. The minority language is often unknown to teachers and other speakers, and the majority language is normally assimilated at the cost of the minority language. This type of bilingualism is also known as folk bilingualism, a term defined by Tosi (1982, in Harding and Riley, 1986, p. 24) as "the condition of ethnic groups within a single state that have to become bilingual involuntarily in order to survive". The Indians and immigrant populations in Brazil and the Spanish heritage speakers[5] in the United States are examples of populations that go through this process, losing or having devalued the minority language they speak at home and assimilating the majority language, that is, the language of the society in which they are immersed. There is, however, currently, a huge concern around this issue, and several bilingual advocates seek the preservation of minority languages, with a tendency towards the maintenance model (García, 2009; Flores, 2016, 2017).

---

[5] Spanish heritage speakers have a family or community connection to Spanish, but in the US they are considered English language learners and normally receive their formal education in English, the majority language spoken in the county they live in. Spanish is the language they inherit from their families and the one mainly used at home (Montrul, 2012).

Elite or prestige bilingualism is a type of additive bilingualism, when a second language is acquired without the risk of the individual losing the first language (Lambert, 1974). In these contexts, students are typically educated in two dominant, prestigious languages. This type of bilingualism is, according to Baker and Prys Jones (1998, p. 15), "planned and purposeful". It is a choice, normally of the family, and the parents are aware of the social and economic benefits of such a model. Here, the second language does not pose a risk to the first, as both are valued at school and in the community. Hélot (2006) adds that in such cases the second language acts as a complementary tool in communication, thinking and learning.

García (2009) questions the nomenclature of the models above, as they are based on the idea that bilingual individuals are a "double monolingual", that is, two perfect monolinguals in one individual. The author questions this diglossic perspective of bilingual education, proposing the adoption of recursive bilingual education and dynamic bilingual education, supported by a heteroglossic view of language[6]. For García, an individual's languages are interactive and complementary. In this view, recursive bilingual education programs seek to rescue ancestral languages, revitalizing them, while in dynamic programs there is the development of an additional language to that of the community.

According to García (2009, p. 313), "bilingual education blends approaches towards education in general with approaches that are specific to language education", supporting both the home and the additional language, but also promoting content learning. There are three language education approaches that are often used in bilingual education: the grammatical approach, the communicative and the cognitive approach, all of which will be detailed below (García, 2009).

The grammatical approach, based on Behaviorism, emphasizes the rules and the structure of the language that is being acquired. The grammar-translation method, the direct method and the audiolingual method are examples of this approach, in which rules are explicitly taught to students. It is, in a way, what was done in language classes in regular schools in Brazil for a long time, with language being taught through grammar only, with no focus on communication or use.

The communicative approach, on the other hand, is based on Constructivism and focuses on the learning that takes place through interaction and experience. The immersion instruction and the integrated content-based instruction (ICB) are learning methods based on the communicative approach. Immersion instruction "promotes the use of language that is slow and simplified, with guarded vocabulary and short sentences" (García, 2009, p. 318) to develop the learning of language and content. It is not necessarily a method that involves the use of two languages, as

---

[6] Languages are seen as interactive, complementary and dependent. They are not in competition. They work together, having different functions. One is not stronger or better than the other.

sometimes the immersion takes place in one language only. On the other hand, integrated language-content instruction (ICB) focuses on developing content alongside language knowledge and skills. Background knowledge and visual aids are key to the planning of a class under this method, which is, in fact, prevalent in many bilingual programs around the world as well as in Brazil (García, 2009; Stoller, 2004).

Finally, the cognitive approach sees learning as social and interactive, but also as involving cognitive processing. It distinguishes between three types of knowledge: "declarative knowledge, what we know; procedural knowledge, what we know how to do; and conditional knowledge, the knowledge of when, why or where to use information and skills" (García, 2009, p. 317). The Cognitive Academic Language Learning Approach - CALLA (O'Malley, 1994) is one of the methods that follows the cognitive approach, combining teaching content and language, but with an emphasis on academic language abilities, using explicit instruction to do so. The development of the learner's metacognitive processes is one of the objectives of the cognitive approach and explicit teaching and modeling of learning strategies and language are characteristics of this approach.

García (2009) claims that, independently of the approach, bilingual education can only be recognized as such if it develops practices that encompass the principles of social justice and social practice. The author states that "it is in the blending of the two, in much of the same way as in the blending of languages and cultures that is a result of bilingual education, that bilingual teachers can be effective" (p. 336). Therefore, the students' learning context cannot be threatening to their identities and it should provide opportunities for different language uses and identities. Besides, learning should be the result of collaborative social practices and be socially built (Vygostky, 1978). García (2017) sums up the issue well by stating that bilingual education is "the use of different linguistic practices to educate" (p. 2); educate accepting and valuing linguistic differences, individual's linguistic repertoire, working to promote bilingualism, or multilingualism, in order to add, not diminish, including the voice of each individual, each student and each teacher in this story.

Most of the studies mentioned above do not account for the reality of our country and might not represent the bilingual education models that have been being developed in Brazil for the past decades. Most literature in the area refers to language communities where two languages are present in the students' lives on a daily basis, which is not the case of most bilingual schools in Brazil. Bilingual education in Brazil will be the topic of the next section, seeking to analyze such practices and models within the Brazilian reality.

2.1.3 Bilingual Education in Brazil

There are a variety of languages that make up the Brazilian linguistic scene, many of them indigenous, immigration, contact languages in border regions, and also LIBRAS, the Brazilian sign language. The new Constitution of the Federative Republic of Brazil from 1988 guarantees indigenous peoples the right to their own languages, cultures and education (Morello, 2012), recognizing the diversity of languages that make up our country.

Some advances have been made over the years regarding the recognition of the huge linguistic diversity that characterizes Brazil. In 2002, bilingual education for the deaf was guaranteed by law. According to Megale (2018), in this educational model, deaf students have the possibility of having instruction through LIBRAS, being Portuguese taught as a second language, guaranteeing the right to an education that takes place in the children's mother tongue.

In addition, since 2002, with the federal decree for the co-officialization of minority languages, some Brazilian municipalities have established municipal co-officialization policies. According to Institute for Research and Development in Linguistic Policy (IPOL), there were eleven co-official minority languages in Brazil by 2017: Tukano, Neengatu, Baniwa, Guarani, Akwê Xerente, Macuxi, Wapichana, Pomerano, Talian, Hunsriqueano and Alemão, distributed in twenty-eight municipalities.

Based on these changes and supported by the law, the right to bilingual education began to be recognized in Brazil for speakers of mother tongues other than Portuguese, being divided into four categories: Indigenous Bilingual Education (Lei de Diretrizes e Bases da Educação Nacional, Law 9,394/1996), Bilingual Education for the Deaf (Law 10,436/2002), Bilingual Education in Immigration Contexts and Bilingual Education in Border Contexts -Border Intercultural Bilingual School Project (PEIBF), currently called the Border Intercultural Schools Program (PEIF),(Action Plan 2006-2010, Mercosul), this one with the main objective of promoting cultural exchange between professors and students from Mercosur countries (Megale, 2018).

In addition to bilingual education targeted to these minority groups, which are guaranteed by law, a specific kind of bilingual education programs have recently emerged in Brazil, mostly involving private schools that serve a clientele that seeks to guarantee the opportunity for their children to learn a prestigious language, such as French, German and, mainly, English. There are also a few public schools that offer bilingual education programs or curricula, but in smaller numbers. According to the latest MEC school census[7], the country has around 40,000 private schools, 21% of the 184,100 Brazilian school units. The Brazilian Association of Bilingual

---

[7] http://portal.mec.gov.br/ultimas-noticias/33541-censo-escolar

Education (ABEBI[8]) estimates that at least 3% of these private schools (1,200) have some sort of bilingual education.

The legislation for bilingual schools is very recent. There are official documents that establish norms for the provision of bilingual education in the states of Rio de Janeiro (2013) and Santa Catarina (2016) in elementary schools. Recently, the National Curriculum Guidelines for Multilingual Education[9] were discussed and formulated (yet to be homologated) in order to organize and define what characterizes a bilingual school in Brazil, differentiating it from schools with an Extended Curriculum in Additional Language and from International Schools.

According to these Guidelines, "Bilingual Schools are characterized by promoting a single, integrated curriculum taught in two languages of instruction, aiming at the development of linguistic and academic competencies and skills of students in these languages." The guidelines also establish rules regarding the workload in both languages, establishing that the bilingual curriculum must be offered to all students in the school, at all levels of education. As for the instruction time in L2, it must offer between 30 and 50% of the curricular activities in Early Childhood Education and Elementary School, and in High School at least 20% (twenty percent) of the workload in the official curriculum needs to be offered in the additional language, with the possibility of also including training itineraries in the additional language as well.

In Brazil, bilingual schools are mainly focused on adding a prestige language to the student's repertoire and do not represent a risk to their mother tongue. It is usually a language that is not spoken at home, being the school the place for learning and practicing. There is also an implicit assumption that learning an additional language does not harm the maintenance and development of the children's mother tongue. Within the contexts of prestige bilingual education, all languages are treated as a source for learning and are not considered a threat to one another.

In the next section, we present a review of models that describe how human speech is produced in L1 and L2.

2.2 SPEAKING IN ONE AND IN TWO, OR MORE, LANGUAGES

This section introduces theoretical models on the production of speech, both in monolinguals (Levelt, 1989) and bilinguals (Green, 1986; De Bot, 1992; Poulisse & Bongaerts, 1994). The present study aims at analyzing the evolution of bilingual students in terms of writing in both languages, but also at accompanying this process in oral languages, in their L1 and L2, looking at the connectedness and syntactic complexity of the students' languages in their oral and

---

[8] Associação Brasileira do Ensino Bilíngue: ABEBI- http://abebi.com.br
[9] http://portal.mec.gov.br/docman/setembro-2020-pdf/156861-pceb002-20/file

written productions. Therefore, it is of paramount importance that we also consider essential aspects of speaking as a human ability.

2.2.1 Speaking in L1

According to Levelt (1999, p. 83), "the ability to speak is one of the basic ingredients of human life". In this section we will characterize speech production in monolinguals, presenting the *blueprint of the speaker*, as described by the author.

Even though we do not know for sure all the stages of evolution regarding speech production, we do know that there have been two landmark developments: the development of supralaryngeal articulation under neo-cortical control, and the emergence of Theory of Mind (ToM[10]). The first landmark involves the evolution of the articulatory system, which is capable of working under intentional control and which is species-specific. With this upgrade, human beings have control over the voice from the larynx area (Levelt, 1999). The second landmark, on the other hand, is attributed to social competence: the emergence of ToM. A larger size of neocortex in man seems to be one of the most significant differences between human brains and other primates. Neocortical areas are dedicated to face and voice recognition, to the recognition of intention (facial expressions) and to the processing of speech. Levelt (1999) explains that ToM allows us to build up complex knowledge structures about our social environment.

The first vocalizations produced by children are those of internal feelings and sensations. They are interjections produced by our vocal organs. Levelt (2019, p. 2927) claims that they are "direct, nonmediated expressions of primary sentiments, such as pain, surprise, disgust and aversion." These sentiments excite speech organs, which in turn, produce sounds. The articulatory system begins to mature when the infant starts to babble, approximately at the age of seven months. In the beginning, babbles are not meaningful. They are articulatory-motor activities that are reinforced by feedback, oral and gestural. Babbling is also present in deaf children of deaf parents, who hand babble in the same period (Petitto & Marentette, 1991). In this stage, the two systems (articulatory and meaning) are still separate.

At around the age of 12 months, links start to be established between the above mentioned systems, which become even more connected during the second year of life. Elbers (1982) has shown that the first spoken words are usually babbles that resemble meaningful words.

Levelt (1999, p. 86) explains that these systems play different roles in speech generation.

10

The semantic/syntactic system is there to map the conceptualization one intends to express onto some linear, relational pattern of lexical items ('lemmas'), a 'surface structure', for short. The function of the phonological/phonetic system is to prepare a pattern of articulatory gestures whose execution can be recognized by an interlocutor as the expression of that surface structure, and hence of the underlying conceptualization. I will call it the 'articulatory score'.

In other words, one's ideas must be structured and organized into items that can actually be externalized and then understood by one's interlocutor. Levelt's blueprint of the speaker (Figure 2) is presented below.

Figure 2 - A blueprint of the speaker
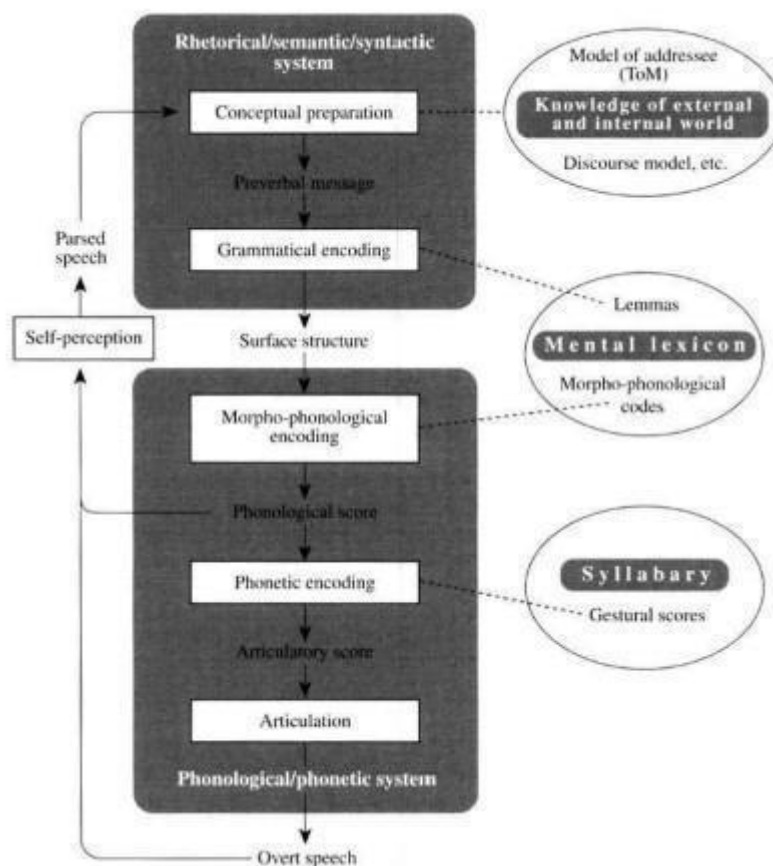
Source: Levelt, 1999, p.87

The blueprint of the speaker is organized as shown in Figure 2. The conceptual preparation is the phase in which each speaker generates a message. This process can be done individually or interactively with the interlocutor. In order to do that. the speaker must access various knowledge sources (diagrammed as ellipses) so as to mind the knowledge shared with the interlocutor, for instance. The message indeed needs a conceptual structure, that is, it must be expressible in words.

The concepts in the message will activate the corresponding syntactic words ('lemmas'). In this phase, the speaker uses the lexical-syntactic information to build the 'surface structure'. Once the lemma is selected, the speaker has access to the item's morphological and phonological composition. Each of the syllables in the phonological score must trigger an articulatory gesture, those gestures that infants began to produce by the end of their first year of age. According to Levelt (199, p.88), "phonetic encoding is the incremental generation of the articulatory score of an utterance". It is executed by the laryngeal and supralaryngeal apparatus, producing the overt speech. This process is monitored by ourselves. As we speak, we monitor our internal and overt speech. It is the same system we use for listening to others.

All these mechanisms work in synchrony. They are simultaneously active, in an overlapped way, like the tiles of a roof. As we speak, we are organizing the next phrase, in a continuous way. Basically, according to Levelt (1999), speech production involves three stages: conceptualization, formulation and articulation. Conceptualization regards the selection and ordering of relevant information and is the phase where the intentions of the speaker are converted into language. It is the preverbal stage, containing the necessary information to convert meaning into language. The formulation phase is the stage in which the preverbal message is converted into a phonetic plan and the selection of the right words takes place; it is also in this phase that grammatical and phonological rules are applied. Finally, articulation is the stage in which the speech plan is converted into actual speech. In simpler words, it is the output of the formulation. These stages are also present in bilingual speech production. In his well-known model of bilingual production, De Bot (1992) includes these components, and the author also discusses it in his more recent multilingual processing model (De Bot, 2004), following the three main stages that have been proposed for monolinguals.

2.2.2 Speaking in L2

Levelt's (1999) blueprint of the speaker was designed for monolingual speakers and needed to be adapted or remodeled in order to account for bilingual speakers. In his adaptation, Green (1986) came up with the idea of 'tags', as labels to be associated with each item. As bilinguals can switch from one language to another, Green proposes that there must be a 'specifier', that is, a device that controls requirements for either speaking L1 or L2. The author says that selection is not only about increasing the activation of one of the languages, but also suppressing the activation of the other one. Figure 3 below presents the scheme of how the system controls the language that will be used for the output.

Figure 3 - An inhibitory control model for a bilingual speaker within the control, activation and resource framework
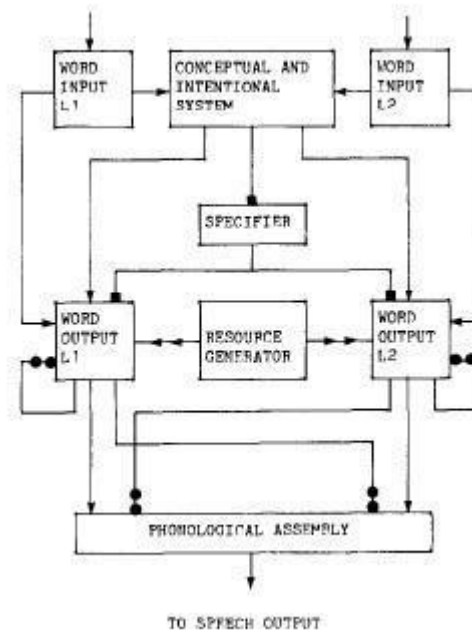
Source: Green (1986, p. 212)

This system is used for regulating spontaneous speech as well as translations, but with distinct inhibitory means. Green explains that a more complex form of regulation is needed in translations, for both language systems are required in this situation. According to the author, 1986, p. 217) "since translation into L1 requires that the speaker does not simply repeat the message in L2, it is proposed that suppression of the output from L2 is achieved internally in the same way as a monolingual speaker might avoid simply repeating a word or phrase just read." For the author, in order to select a word, it is necessary that its activation exceeds that of any competitors, that is, it is necessary to regulate the amount of activation.

Words that are less predictable are at a lower level of activation, which makes the speakers have a longer pause before uttering the word. It is important to highlight, however, that knowing two languages cannot be linked to delay or interference only. According to Green (1986, p. 215), being bilingual allows speakers to "output whichever expression first achieves threshold. Hence, code switching need not involve dysfluency." Levels of activation may vary, according to language use, depending on how much the language is spoken and heard, falling when languages are not used.

In a nutshell, when a bilingual speaker wants to speak, a specific language must be selected and the output of the other language must be inhibited. This control of which language to select and which to inhibit is what Green calls the inhibitory control model.

De Bot (1992, p. 9) attempts to adapt Levelt's 'Speaking Model' for the bilingual speaker, intending to change it as little as possible. He explains that "knowledge of the two languages may be represented and stored separately for each language or in a shared system" depending on the linguistic distance between the languages and on how proficient the speaker is in using them. With regard to the speaker's level of proficiency, De Bot (1992) clarifies that the L1 system is flexible enough to include an additional register, but only when the speaker's L2 knowledge is still very limited.

De Bot (1992) adopts Paradis' (1987) subset hypothesis, which "assumes the use of a single storage where links between elements are strengthened through continued use" (De Bot, 1992, p. 11). Lexical items from the L1 and L2 compose separate subsets that are activated to varied extents according to the language currently being spoken (De Bot, 1992; Poulisse, 1997; Weissheimer, 2007), similarly to the proposal by Green (1986).

With respect to the three speech production stages proposed by Levelt, De Bot suggests that there should be some modifications in the conceptualizer. Assuming that it is language-specific, the two production phases, the macroplanning and the microplanning, would have a different behavior: the macroplanning would not be language-specific, whereas the microplanning would be. According to De Bot (1991, p.21), "communicative intentions are given form in the preverbal message, which contains information about the language in which (part of) an utterance is to be produced". The language-specific formulator is then activated, converting the preverbal message into a speech plan, similar to what is proposed in the model put forward by Levelt. The speech plan is submitted to the articulator - which is not language-specific and stores sounds and patterns of different languages.

Along the same lines, Poulisse and Bongaerts (1994) suggest that bilinguals are able to separate or blend the language systems by specifying the language of choice in the preverbal message (Poulisse, 1997; Weissheimer, 2007). Poulisse and Bongaerts' (1994) model, like Green's (1986) and De Bot's (1992), assume that the L1 and L2 systems may share the same mental lexicon and that the words from the two systems carry identification tags (Poulisse, 1997).

These models, though based on Levelt (1986), show that bilinguals and monolinguals have a different *modus operandi*, when considering oral production. The fact of having to select, suppress and search for the appropriate lexical items may have a cognitive demand, especially at early stages of proficiency. As proficiency develops, oral communication tends to become more fluid and fluent. In fact, the speed of the processes behind speech production in L2 (Kirk, 2014) is one of the main difficulties L2 learners encounter, as they have to select among alternatives before they speak, which presents time constraints. Achieving fluency in an L2, which is, according to Lennon (2000, p. 26) "the rapid, smooth, accurate, lucid, and efficient translation of

thought or communicative intention into language under the temporal constraints of on-line processing", imposes a great cognitive demand in L2 learners. Again, it is paramount to understand that in cases where L2 oral productions are not as fluid as L1's, it might be the case that proficiency is permeating this relationship.

In the next section, we present research on the development of writing skills in one and two or more languages, as the objective of this dissertation is to investigate the effects of schooling in two languages on the written production of students. We also review some of the studies in the area, as they may contribute to our analysis.

2.3 WRITING IN ONE AND IN TWO - OR MORE - LANGUAGES

Whether in a first or a second language, writing has always been seen as a skill that requires great effort in order to be acquired (Vasylets, 2021). Also, a skill "that is developed to immensely higher levels in some people than in others" (Bereiter; Scardamalia, 1987, p. 4). What accounts for these individual differences has been the topic of a number of studies that analyze the role of language knowledge and expertise, as well as capacity and efficiency of a writer's cognitive resources (Olive, 2012) and working memory[11] (WM) (Hayes, 2012; Kellog, 1996), for instance.

Several cognitive writing models have been proposed since the 80's. The first ones, though, were designed for adults (Hayes et al., 1987; Hayes, 1996; Kellog et al., 1983; 1985), with the aim of circumscribing the processes, knowledge and types of processing required for the production of a text (Barcellos, 2021).

Writing models based on cognitive processes can be divided into two main categories: the early models (1980s) and the task-centered models (1990s).

During the 1980s, composition researchers studied how cognitive processes interacted during writing (Becker, 2006). Starting with the model proposed by Flower and Hayes (1980), an emphasis in the reviewing process was the main focus of different researchers.

Flower and Hayes' initial model (1980) was divided into three processes: *planning*, *translating* and *reviewing*. In 1981, the authors included a monitor function that allows access to the writer's long term memory. Furthermore, they subdivided the reviewing process in "1)

---

[11] "The term working memory refers to a limited-capacity cognitive system involved in the manipulation and maintenance in active attention of the task-relevant information and the inhibition of task-irrelevant information." (Baddeley, 2003, p.1). Vasylets and Marín (2021) investigate the effects of working memory on L2 writing performance.
"Holding information in mind and mentally working with it (e.g., relating one thing to another, using information to solve a problem)" (Diamond, 2012, p. 137).

evaluation, which provided for specific appraisal of the written text, and 2) revision, which referred to the actual changes" (Becker, 2006, p. 26).

Bereiter and Scardamalia (1983, 1985) developed a compare, diagnose and operate (CDO) planning stage, expanding the evaluation and revising processes initiated by Flower and Hayes (1981). The authors theorized that writers first "compare" their mental text to their written production and, if they see a problem, they "diagnose" what needs to be done to improve their writing. The same authors also investigated the role of advanced planning in increasing students' reflective thinking.

These early models have had an important contribution in shifting the focus so that more emphasis was devoted to the reviewing process. Another considerate shift was seen in the 1990s, with the task-centered models focusing on furthering in-depth analysis of working memory and long-term memory in writing proficiency, as well as on the social and motivational aspects of the writing process (Kellog, 1996; Hayes, 1996; Huub van der Bergh, 1994; 1999). These studies have made it clear that "well-developed reading ability and extensive writing experience expert writers possess expands working memory capacity and long-term memory knowledge" (Becker, 2006, p. 34).

Now we shift the focus to research regarding children. The models that deal with the evolution of the process of writing in beginning writers, who possess less experience, are known as developmental models. They intend to show the evolution in the writing process in writers who are still developing both the graphomotor system and initial spelling issues (Barcellos, 2021). Some of these models will be presented in the next section.

2.3.1 Models of writing for beginner writers

The main theoretical models of writing production were initially designed considering experienced writers. With time, however, researchers realized the need for analyzing the writing process from the perspective of inexperienced writers. In this section, we present some of the models that have focused on the writing development of children.

Bereiter and Scardamalia (1987) proposed a distinction between more and less experienced writers based on their writing strategies: *knowledge-telling* and *knowledge-transforming*. In the knowledge-telling model, as soon as the memories are retrieved from long-term memory[12], they are put into words written in a text. Planning and writing are done together, sentence by sentence, without reformulation or monitoring (Barcellos, 2021), and this might be the procedure followed

---

[12] Long-term memory consists of the repertoire of information that is coded and stored over an extended period of time in the brain (Gazzaniga, 2009).

by less experienced writers. On the other hand, in a knowledge-transforming model, there is constant monitoring of what has been written and the extent to which the written words represent the writer's intentions and ideas. This model is assumed to be engaged by more experienced writers, which is why it is more recurrent with adolescents and adults, in comparison to children, who have not had enough time to improve their writing skills. In addition, this model allows the writer to associate the development of the information to the development of the text.

'Expertise' is an important concept associated with these models and requires time of exhibition and systematic use of the specific ability in order to be developed (Chi, 2006). As the abilities of writing and spelling demand a great deal of cognitive resources, they need to be available for the writers so that they are able to monitor their text production. According to Kellogg (2002), the development of writing *expertise* can be characterized by the effective use of monitoring mechanisms combined with long-term memory resources. It is important to highlight that, during the early years of schooling, handwriting skills and orthography constitute great cognitive demands, which might present difficulties to the flow of the narrative. Until graphomotor transcription becomes automated, there is little availability of attentional resources for the child to monitor or evaluate what is being written. The attentional demands involved in writing seem to be similar to what happens in the beginning of the decoding process when children are first learning to read. As children become so involved in decoding the words, letter by letter, they might initially have difficulty understanding what they have just read, a difficulty that disappears with reading practice.

Another model was proposed by Berninger and Swanson (1994), based on Hayes and Flower (1980), but with adjustments in order to adapt it for children. The model presents three processes: *planning*, *formulation* (or *translating*) and *reviewing*. Planning covers the generation and organization of text content, including all processes involved in the organization of the ideas that will be conveyed in the text. The formulation (or translating) stage is when the sentences are generated in order to convey the ideas selected in the planning stage. Finally, the reviewing stage focuses on evaluating what has been written so far.

The formulation (or translating) stage was later subdivided into two main parts: text generation, which involves the process of transforming ideas in linguistic structures, and transcription, which regards the representation of written symbols and allows for the grammatical and orthographic codification. As the transcription process is very intense, it ends up reducing the WM capacity, which explains why oral stories told by children are more complete than written stories. The simple fact of transcribing what has been said orally to the written media requires attentional resources that may interfere with the production. According to this model, when the

transcription process becomes more fluent and efficient, text generation starts evolving from simple words to sentences and texts.

Berninger et al. (2006) developed the Not-so-simple view of writing, based on the studies of Neuroscience. According to these authors, there is an interaction among transcription, text generation and executive functions[13]. In this model, transcription refers to the graphomotor processes involved in writing, text generation refers to linguistic processes involved in the production of a text, and the executive functions regulate the goals, planning and reviewing processes.

Figure 4 - Not-so-simple view of writing
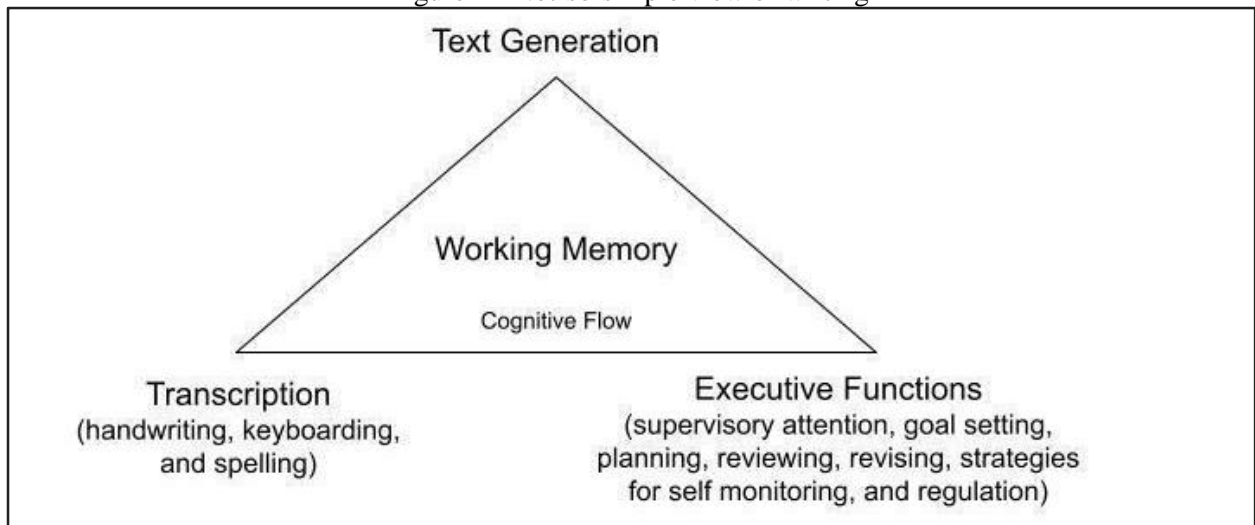
Source: Berninger and Winn (2006).

The transcription phase, which involves "putting words onto paper", that is, transcribing oral material into concrete, written words, requires a lot from the WM capacity, as has been shown before. However, this is not the only phase that demands resources from the WM in beginner writers, which requires them to administer these processes in order to write. Bereiter and Scardamalia (1997) *knowledge-telling* strategy seems to be a good way to diminish the load on the WM. Once the transcription becomes more automatized, children may allocate WM resources to other sub processes of writing. Berninger et al. (1992) showed that children can readily express their ideas orally, but do not, in the beginning, possess the knowledge necessary to orthographically represent what is expressed in oral language.

---

[13] "Executive functions (EFs; also called executive control or cognitive control) refer to a family of top-down mental processes needed when you have to concentrate and pay attention, when going on automatic or relying on instinct or intuition would be ill-advised, insufficient, or impossible." (Diamond, 2012, p. 136).

Text generation, on the other hand, is related to the linguistic processes required for the production of a text: words, sentences and discourse. In this sense, it is believed that well developed oral abilities contribute to the development of text generation (Barcellos, 2021).

The executive functions (EFs) encompass the cognitive basis of writing and include complex abilities, such as working memory mechanisms and high-order executive functions, which require the combined functioning of the central executive functions, and are responsible for the goal setting, planning and review.

Summing up, even though the models differ in many aspects, they all seem to account for the cognitive demand required for transcription, especially in beginner writers, when this process is not totally automatized. They also emphasize the difference between oral and written texts in terms of cognitive demands, explaining the advantage oral texts have over written texts, since the demand for transcription requires resources from the WM.

The models presented above are not specifically designed for bilingual writers, but writing *per se*. If we consider writing in an L2, we have to account for the multiple skills that are involved in the process, such as linguistic knowledge, knowledge of genre, metacognitive knowledge of the writing process, and fluency (Schoonem et al., 2011; Vasylets, 2021; Weigle, 2005). In fact, writing in an L2 may add some difficulties to the writing process due to the possible linguistic gaps in L2 and the lack of automation of access to L2 linguistic representations.

Considering the cognitive demands involved in writing itself, one can imagine that writing in an L2 is an even more challenging process. Having this in mind, in the next section we present previous studies that have investigated the process of writing in an L2 in order to contribute to the discussion on this topic.

2.3.2 Writing development in two languages

Writing in two languages has been seen as something way too complicated than it really is. Lack of research on the subject made people believe that having this process in two languages would cause more harm than good. With this in mind, many children were "forced" to go through this process counting only on part of their repertoire.

More up-to-date research has shown the opposite: the positive effects of having children lean on their full repertoire when learning how to write. Cummins' *Underlying Linguistic Competence Hypothesis* (1979) has backed this idea, positing that literacy skills in a first and second (or additional) language influence each other through an underlying linguistic competence (Cummins, 1979). Following this perspective, it is possible to assume that if one language affects the other, leaving one of them behind will definitely impact the overall outcome.

Following the same lines, Dworin (2003) reinforces that development in one language has a direct effect on the development of the other language. Other researchers have also argued that "bilingual individuals draw on all their linguistic and cultural resources as they learn how to read and write" (Butvlofski et al., p. 2).

On the same lines, Groff and Bellamy (2020, p. 2) argue that "the development of receptive and productive skills in both speaking and writing are all interrelated in contexts of biliteracy, and the development of biliteracy skills is facilitated when children are encouraged to draw on their linguistic resources." Corroborating with this data, S. Montanari et al. (2016, p. 54) conducted research with bilingual students in an Italian/English immersion program, stating that "students were able to develop moderate-to-advanced writing skills in English, even if primarily instructed in Italian". These results are in line with studies with other combinations of languages, such as Spanish-English (Howard et al., 2004), Korean-English (Bae, 2007) and Mandarin-English (Padilla et al., 2013), indicating the influence of L1 and L2 writing skills in typological and orthographically related languages (Montanari, 2016).

Besides, Montanari et al. (2016) reinforced the correlation between advanced writing ability in L1 and L2, as well as limited writing ability in L1 and in L2, which, according to the authors, is possibly related to the notion of mutual influence of L1 and L2 writing skills.

Howard et al. (2004) conducted a longitudinal study with two-way-immersion students from third to fifth year over a three-year-period, looking at writing, reading and oral proficiency, examining the differences and growth patterns for native Spanish and nonnative English speakers. The narrative ability specifically involved the collection of three narratives per year, in both languages, Spanish and English. The results indicate that "there is clearly a relationship between writing ability in English and Spanish for both groups of students, and that relationship seems to be fairly consistent both over time within each language group and across language groups" (p. 19). The authors also highlight that there was an initial gap in performance in the third year, having native speakers outperformed nonnative speakers in their writing assessment. However, over the three-year-period, the gap was closed, as Spanish native speakers demonstrated more growth in English and English native speakers demonstrated more growth in Spanish. One of the concerns is that native English speakers had a certain advantage in the development of literacy skills, which might be related to socioeconomic factors.

One of the difficulties faced when working with writing in two or more languages is the fact that, usually, languages are kept apart, as if they did not influence each other and the students overall performance. Hornberger (2003, p. 26) talks about it as a continua, the continua of biliteracy, explaining that "the more the contexts of their learning allow them to draw on all points of the continua, the greater are the chances of their full biliterate development."

Sparrow et al. (2014) conducted a study showing the development of bilingual children's writing. Their hypothesis was that

if students were progressing along a trajectory toward biliteracy, their Spanish literacy would be slightly more advanced that their English literacy, but a large discrepancy would not appear between the two, and students would be considered to be on a positive trajectory toward biliteracy (p. 160).

The authors concluded that "the best way to understand emerging bilingual children's biliterate development is to examine what they can do in both languages, as when looking at each language independently, one risks underestimating their full capacities" (p.164). They argue that it is difficult to look at students' writing samples isolatedly, as they do not represent all their weaknesses, nor all their strengths.

Another study by Sparrow et al. (2014) supports biliteracy and the importance of explicitly teaching students to make connections across languages. The authors believe that these connections "do not confuse them or hinder their academic development in either language" (p. 40).

Again, the studies conducted so far do not represent the reality of bilingual schools in Brazil. In fact, there is not an agreement on when to start the process of writing/reading in the L2, maybe due to the lack of research representing our students' situations. The lack of scientific evidence may lead to fear and doubt. Schools and teachers must focus on studying biliteracy in order to show families that the process is safe.

It is imperative that teachers understand how students use languages as they develop biliteracy for there is no predictable path toward biliteracy. Each person's trajectory is different, unique. Thus, the importance of examining each one's development from a bilingual perspective.

Even tough there has been an increase in biliteracy research, much more needs to be done in order to expand knowledge of the processes that contribute to children's early biliteracy development (Reyes, 2012, p. 323). It is critical that we reexamine biliteracy practices at school so as to better help teachers and students during this path, which is one of the aims of this project.

In the next section, we will discuss Graph Theory as a tool for investigating language development and describe SpeechGraphs, the specific tool used in this research. Other studies in the area will be presented to corroborate the idea.

2.3 GRAPH THEORY

Graph Theory is a branch of Mathematics that deals with the relationship between elements of a set (Luz, 2018). The study of graphs began in 1736, when the Swiss mathematician Leonhard Euler solved the enigma of the Konigsberg Bridge Problem. This city had two islands, connected to the mainland, and seven bridges. Residents wondered if there was a way to go over all the bridges and return to the starting point. The mathematician solved the puzzle by transforming the paths of bridges into lines and their intersections into points, thus creating a graph (Alexanderson, 2006; Da Luz, 2018). However, the answer was negative, because in order to pass through all the bridges and return to the starting point, there should be an equal number of bridges entering and leaving the city, that is, an even number, not an odd number of bridges. Since then, graph analysis has been used in a range of scientific studies in view of its adaptability and correspondence to different areas of knowledge.

Figure 5 - Graph of the bridges of Königsberg

Source: Wikipedia[14]

Based on Graph Theory, researchers from the Federal University of Rio Grande do Norte created a software called *SpeechGraphs*, which represents the relationships between words in a oral or written text through graphs, revealing characteristic patterns of discursive connectivity of different groups or enunciators (Mota et al., 2014). According to Mota et al., (2012), a graph represents a network with nodes linked by edges. It is composed of a set of points called *nodes* (N) and of lines connecting them, called *edges* (E). The nodes are the representation of the words, while the connections or relationships between the represented elements are established by the edges.

Mota et al. (2017) explain that "by representing each word as a node and the temporal sequence of consecutive words as directed edges, it is possible to calculate attributes that characterize graph structure" (p. 1). Selecting the pertinent attributes for each analysis is an

---

[14] https://en.wikipedia.org/wiki/Seven_Bridges_of_K%C3%B6nigsberg

important part of the work with graphs. In the table below, all the elements that are possible to be analyzed with the use of *SpeechGraphs* are described.

Table 1 - Mathematical definition and psychological interpretation of SpeechGraph attributes (SGA)

Table 1

Mathematical Definition and Psychological Interpretation of Speech Graph Attributes (SGA)

SGA	Mathematical definition	Psychological interpretation
N (nodes)	Number of nodes	Number of different words, measures lexical diversity
E (edges)	Number of edges	Number of links between words
RE (repeated edges)	Sum of all edges linking the same pair of nodes	Number of links between two words; measures recurrence
PE (parallel edges)	Sum of all parallel edges linking the same pair of nodes given that the source node of an edge is the target node of the parallel edge	Number of links between two words with opposite directions; measures recurrence
L1 (loop of one node)	Sum of all edges linking a node with itself, calculated as the trace of the adjacency matrix	Numbers of repetitions of the same word in sequence; measures recurrence
L2 (loop of two nodes)	Sum of all loops containing two nodes, calculated by the trace of the squared adjacency matrix divided by two	Number of sequences of two different words; measures recurrence
L3 (loop of three nodes)	Sum of all loops containing three nodes (triangles), calculated by the trace of the cubed adjacency matrix divided by three	Number of sequences of three different words; measures recurrence
LCC (largest connected component)	Number of nodes in the maximal subgraph in which all pairs of nodes are reachable from one another in the underlying undirected subgraph	Number of different words in the largest component in which all the words are connected by a path of edges; measures how well connected the words of the report are
LSC (largest strongly connected component)	Number of nodes in the maximal subgraph in which all pairs of nodes are reachable from one another in the directed subgraph (node a reaches node b, and b reaches a)	Number of different words in the largest component in which all the words are mutually connected by a path of edges; measures how well connected the words of the report are
ATD (average total degree)	Given a node n, the total degree is the sum of "in and out" edges. Average total degree is the sum of total degree of all nodes divided by the number of nodes	Given the word X, total degree is how many links this word has with any other words in the report. ATD is the average total degree of all words in the report
Density	Number of edges divided by possible edges ($D = 2 \times E / N \times (N-1)$), where E is the number of edges and N is the number of nodes	Number of direct word links divided by all the possible word links (using all the different words in the report)
Diameter	Length of the longest shortest path between the node pairs of a network	Length (in words) of the path linking the most distant pair of words in the report
ASP (average shortest path)	Average length of the shortest path between pairs of nodes of a network	Average of all the shortest paths between every pair of words in the report.
CC (average clustering coefficient)	Given a node n, the clustering coefficient map (CCMap) is the set of fractions of all n neighbors that are also neighbors of each other. Average CC is the sum of the clustering coefficients of all nodes in the CCMap divided by number of elements in the CCMap	Given the word X, CC of X is a measure of how many words directly linked to word X are also directly linked to each other. The average CC is the average CC of all different words on the report

Source: Mota et al. (2016)

As the table shows, besides nodes (N) and edges (E), graph attributes also include Repeated Edges (RE), that is, the number of links between two words; Parallel Edges (PE), number of links between two words, but with opposite directions; Loops of nodes (L), which are the number of repetitions of the same word in order, L1, L2, L3. In addition, the software allows for the analysis of global attributes such as Average Total Degree (ATD), showing how many links a word has with the other words in the text; Density (D), representing the direct word links divided by all the possible word links; Diameter (D), measuring the length between the most distant pair of words; Average Shortest Path (ASP), measuring the mean of all the shortest paths between pairs of words; and Average Clustering Coefficient (CC), which measures how many words directed to word X are also linked to each other.

In the studies reported in the present dissertation, the attributes that will be most used are LCC and LSC. Mota et al. (2017) define the LCC as the "largest set of nodes directly or indirectly linked by some path" (p. 2) and the LSC as the "largest set of nodes directly or indirectly linked by reciprocal paths, so that all the nodes in the component are mutually reachable" (p. 2). With regard to LSC, Mota et al. (2019) explain that "when a word is repeated, it closes a cycle in which any two participating nodes are mutually reachable; the LSC corresponds to a special kind of long-range word repetition in which the cycle closed is the largest within the graph assessed" (p. 2). Both LSC and LCC attributes measure how well connected the words of the report are, but the LSC appears to be a more reliable connectedness attribute as it closes a larger cycle of repetition of words (Weissheimer; Costa, 2021).

Connected speech can be understood as textual cohesion, representing the unity of a text. Malcorra *et al* (2021, p. 906) argue that "connected speech is a complex task that recruits various linguistic systems and cognitive domains (such as executive functions and memory)", which makes it an important measure in the analyses of language and cognitive impairments. In fact, deficits in semantic, episodic and working memories are related to speech-connectedness. An easy and non-invasive way to measure speech connectedness is through the representation of graphs, which have already been used and proved to be able to differentiate oral narrative productions of healthy adults from Alzheimer's Disease patients (Malcorra *et al*, 2021). Figure 6 below shows the representation of some graph attributes.

Figure 6 - SpeechGraph attributes
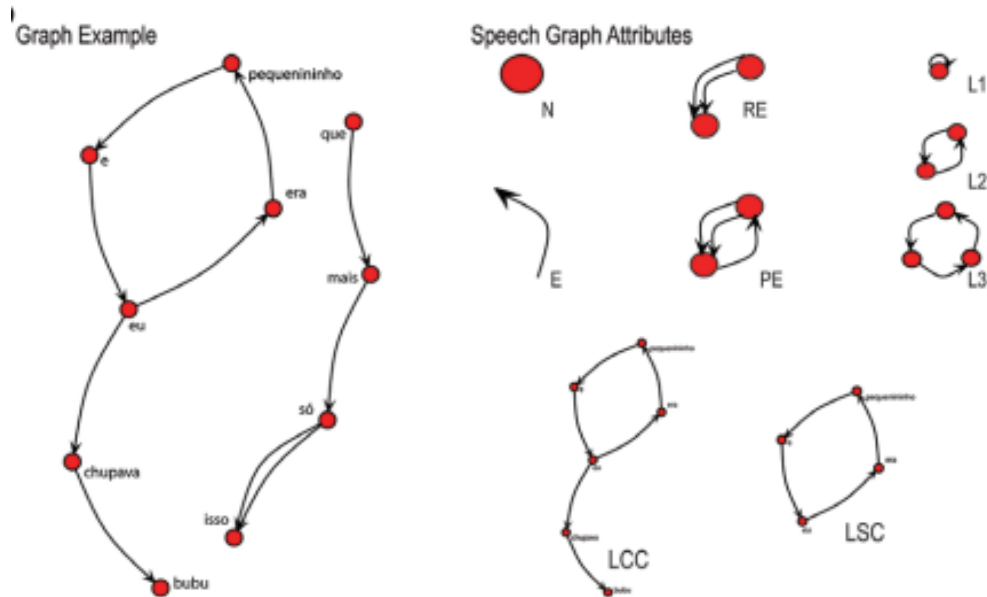
Source: Mota et al. (2014).

In the example above, in the sentence "Eu era pequenininho e eu chupava bubu", the words are connected in a larger component, the LCC, being linked by the word EU, which closes a cycle and connects it to the other part, sharing the same word. The number of nodes produced reflect a larger vocabulary, which is associated with a greater capacity to store and retrieve information. As Mota et al. (2014) point out, a report, in order to be meaningful, needs to have solid links among the events reported. One of the ways to measure these links is the number of nodes in the LCC and LSC. According to the authors, less word recurrence (RE), on the other hand, is associated with higher IQ, ToM, and school performance, reflecting a more developed working memory, without the need of repeating words that are unnecessary.

As for their categorization, graphs can be directed or undirected. Undirected graphs show links between node-represented elements, but without showing their direction. Directed graphs, on the contrary, present arrows instead of simple lines, pointing to the direction they appear, showing which nodes establish a connection with other nodes and which ones do not establish this connection. The following graph (Figure 7) has twelve nodes connected by 12 directed edges[15], represented as arrows, whereas the next one (Figure 8) has twelve nodes, connected by 12 undirected edges, represented as lines.

---

15 All graphs were created using the SpeechGraphs software.

Figure 7 - A directed graph
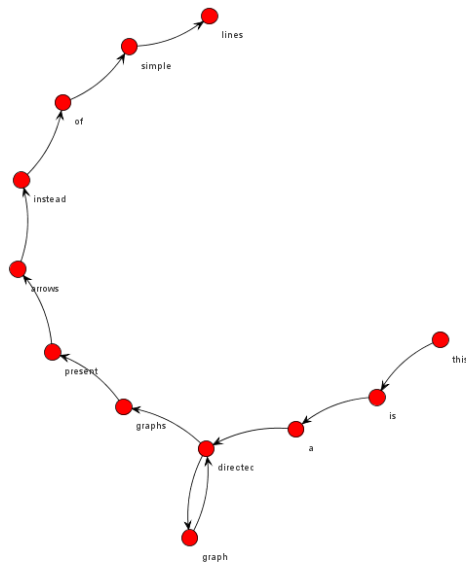
Figure 8 - An undirected Graph
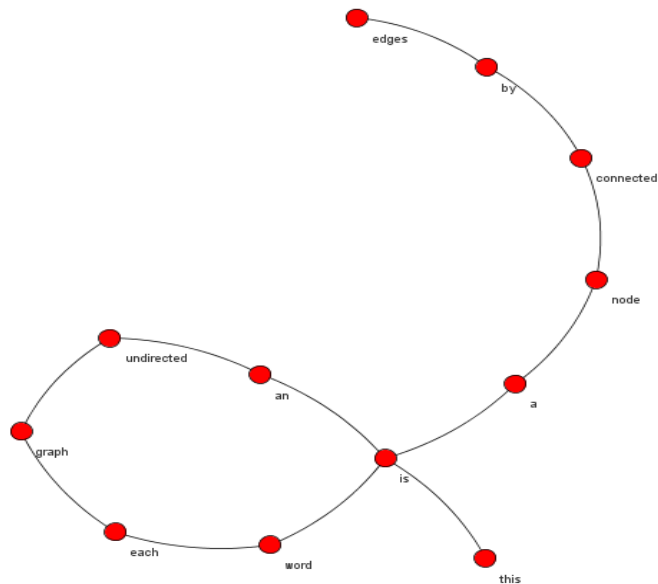
Directed graphs tend to be more informative, since they allow us to visualize and analyze the directions in which words are being connected – or not. Undirected graphs, on the other hand, might be more confusing, as one does not have the indications of what path to follow. In figure 3, the reader knows where to start reading and how to continue.

In recent years, Graph Theory has been applied in studies on Neuroscience (Mota et al., 2012; 2014), Education (Mota et al., 2016; 2019) and Psycholinguistics (Luz, 2018; Leandro, 2021; Lemke *et al.*, 2021). Graph Analysis was used in pioneering work to assist in the diagnosis of thought disorder in adult manic patients (Mota et al., 2012). The participants were asked to talk about a dream and their speech was analyzed using Graph Theory. The researchers were mainly interested in verifying if connectedness and recurrence levels were different in typical participants compared to participants with Alzheimer Disease. Results showed that alterations in thought connectedness in psychotic patients can be measured through graph attributes.

Malcorra et al. (2021) organized a research study with two groups of elderly participants: a control group with typical development, and a group with patients diagnosed with Alzheimer's Disease (AD). They were all asked to talk about a sequence of seven images presented to them. This task was intended to be more naturalistic and representative of language use in daily life than the production of isolated words, used in verbal fluency tasks. As expected, the AD group has less connected narratives than the control group.

Graph Theory has also been applied in Education research (Mota et al., 2016; Da Luz, 2018; Leandro, 2021), proving to be a simple, fast, and noninvasive evaluation of speech connectedness, associated with cognitive and scholastic performance, which may help to develop better intervention strategies in naturalistic settings. In fact, Mota et al. (2016) hypothesized that if connectedness decreased in people with cognitive deficit, it might increase in typical children as they advance in their school trajectory (Weissheimer et al., 2021), starting a series of studies with school children and typical subjects using SpeechGraphs.

In one of the studies, Mota et al. (2016) longitudinally followed a group of second and third year students from public schools using graph theory to analyze student performance. The authors collected autobiographical memory reports from different periods, a dream report, as well as an account from three affective images (one positive, one negative and one neutral). Theory of Mind (ToM[16]) tests were also applied: the Sally-Anne task, and three versions of the Picture Sequence Test (PST; Baron-Cohen; Leslie; Frith, 1986), totaling four skill tests of ToM.

---

[16] "The concept of "Theory of Mind" (ToM) is defined as the cognitive ability to represent one's own and others' mental states, in terms of thoughts and beliefs, but also of desires, demands and feelings, so that one can explain and predict behavior." (Perrotta, 2020, p. 1).

Subsequently, the Raven's Progressive Matrices Test (Angelini et al., 1999; Raven, 1936) was used as a measure of IQ and, finally, the results of the participants in Provinha Brasil[17] were used. In the analysis, the memory reports, as well as the PST test reports, were transcribed and analyzed using the SpeechGraphs tool. The results showed that graph connectedness measures were able to explain children's reading performance and development, even when IQ and ToM measures were controlled for.

In another study, Mota et al. (2019) investigated whether LSC (one of the SpeechGraphs connectedness attributes) would be a good predictor for memory capacity and whether such a relationship could be correlated with reading. In this study, verbal reports of images shown to children were collected as a measure of verbal short-term memory. A few weeks later, the Raven's Progressive Matrices Test, a non-verbal measure of IQ, was also applied, and the children's reading fluency was assessed through a Words and Pseudowords task. New oral reports were collected one year after the first collection, in addition to four working memory tasks (Pearson's Automated Working Memory Assessment - AWMA). As in other studies, the SpeechGraphs tool was used in the analysis of oral data, and a relationship was found between speech connectedness and short-term memory.

Having been used with oral reports before, Luz (2018) used graph analysis to investigate connectedness patterns, but this time in the written texts of 181 children, divided into groups of good readers, bad readers and dyslexics. The task required the children to produce a written story based on comic book images, with no restrictions on time or length. Graph attributes ranked good, bad, and dyslexic readers, revealing patterns of textual connection, measured by the number of nodes and edges, LCC, density, and ASP. The author argues that the theory of graphs represents a new methodological framework in the assessment of reading fluency.

Another relevant study that adopted Graph analysis was the one conducted by Leandro (2020), who examined patterns in the oral production of two groups of adult L2 English speakers, one group formed by pre-intermediate level learners and the other by proficient bilinguals. The author analyzed the oral production of 68 volunteers in terms of speech measures and graph attributes. In general, the more fluent the participants, the more connected their texts were and fewer repetitions they produced. The author claims that in this study Graph Analysis successfully explained fluency in L2. Like Luz (2018), Leandro (2021) argues that graph analysis can be used

---

[17] It is a federal diagnostic assessment of literacy skills in Portuguese and in Mathematics taken by children enrolled in the second year of elementary school in public schools in Brazil. The exam takes place in the beginning and end of the school year to allow the diagnosis and assessment of the evolution of student learning. https://www.gov.br/inep/pt-br

as an alternative for the evaluation of production in L2, since the analysis with this tool proved to be effective in accurately distinguishing levels of proficiency of bilinguals.

Finally, Weissheimer and Costa (2021) investigated the effects of simultaneous biliteracy on the levels of connectedness and reading fluency, in English and Portuguese, with a group of fifth graders enrolled in a bilingual school in the Northeast of Brazil. Results show that reading fluency and speech connectedness develop in parallel in both languages. Children who are more fluent readers in their L1 (Portuguese) are also more fluent in their L2 (English). Also, children who produced a more connected speech in their L1 are also the ones who show more connectedness in their L2.

As indicated in the studies above, graph analysis has been shown to be an important and efficient tool to be used in research that investigates connectedness in oral and written production in naturalistic settings, not only in clinical contexts, but also in educational ones, proving to be a fast and simple tool to be used. In the next section, we will discuss syntactic complexity and its role in assessing language production and development in L2.

2.4 SYNTACTIC COMPLEXITY

Syntactic complexity has been seen as an important construct of analysis for language assessment, especially in the L2 field (Hawkins, 2001; Lu, 2010) and over the years it has been used to estimate development based on people's speech or texts. Hawkings (2001) argues that syntactic complexity develops over time, with some features developing earlier than others. As we will see below, syntactic complexity has also been associated with language proficiency. According to Crossley and McNamara (2014), "syntactic complexity refers to the sophistication of syntactic forms produced by a speaker or writer and the range of variety of syntactic forms produced" (p. 67). The authors claim that

> researchers have focused on L2 syntactic development under the notion that the ability to arrange words syntactically into phrases and phrases into clauses demonstrates the capacity to manipulate a language's combination properties, which is argued to be a strong indicator of general language acquisition (Crossley; McNamara, 2014, p. 66).

One of the most traditional methods of measuring syntactic complexity is through the counting of T-Units, a term introduced by Hunt (1965). In fact, his initial idea was to call them "minimal sentences", like a simple sentence. However, in order not to have any confusion due to the different meanings of the word *sentence*, Hunt decided to call them "Minimal Terminal Units", as "they would be minimal as to length, and each would be grammatically capable of being

terminated with a capital letter and a period" (p. 37). To make it simpler, the Minimal Terminal Unit turned into an M T Unit, which ended up being nicknamed T-Unit. At first, this measure was designed to assess syntactic complexity in L1. Only later was it used to measure syntactic complexity in L2 and started to be associated with L2 proficiency.

Simply put, a T-Unit is "one main clause with all the subordinate clauses attached to it" (Hunt, 1965, p. 36). The following sample (Hunt, 1965, p.21), punctuated as one sentence, 11 clauses, 6 T-Units, 68 words-long, is organized into T-Units.

1. I like the movie / we say about Moby Dick, the white whale.
2. The captain said/ if you can kill the white whale, Moby Dick, / I will give this gold to the one / that can do it.
3. And it is worth sixteen dollars.
4. They tried and tried.
5. But / while they were trying/ they killed a whale and used the oil for the lamps.
6. They almost caught the white whale.

As shown in the example above, several T-Units are made up of single clauses, while others are multi-clauses. Number 2, for instance, is formed by a main clause, with an embedded noun clause, plus an adverbial *if* clause and an adjective clause. Numbers 1 and 5 are organized with two clauses each.

According to Hunt (1965, p. 38), "the T-Unit has the advantage of preserving all the subordination achieved by a student, and all of his coordination between words and phrases and subordinate clauses". Based on the T-Unit, a new ratio was introduced so as to better calculate subordination.

Subordination Index= <u>Number of all clauses</u>

Number of T-Units

In the example above, the Subordination Index (SI), also known as clausal density (Scott, 1988; Guttieres-Clellen; Hofstetter, 1994; Scott; Stokes, 1995), Frizelle et al., 2018), would be calculated as follows: 11/6= 1,83. There are 6 T-units, being numbers 1, 2 and 5 made up of more than one clause (number one, 2 clauses; number two, 4 clauses, and number five, 2 clauses) added to the other clauses, making a total of 11 clauses. Hence the total number of clauses is divided by the number of T-units.

Hunt (1970) used the concept of T-Units as an intermediary structure between the clause and the sentence, using five measures:

mean word per sentence (w/s)

mean T-unit per sentence (T/s)

mean words per T-unit (w/T

mean clauses per T-unit (c/T)

mean words per clause (w/c)

In his thesis, Hunt proposes that as children mature mentally, they tend to embed more of their elementary sentences, that is, they tend to use more subordinate clauses and this is a psychological phenomenon. Hunt argues that "as the mind matures it organizes information more intricately and so can produce and receive more intricately organized sentences" (Hunt, 1970, p. 58).

Sparrow et al. (2014) used T-Units to analyze writing samples of 25 bilingual students over a three-year-period. In their research, they used the number of words per T-Unit to measure students' growth over time, illustrating "students' ability to produce longer and more complex written sentences" (Lanauze; Snow, 1989, p. 327, in Sparrow et al., 2014, p. 162). The authors compared students' writing complexity in Spanish to their writing in English and also compared each individual student's writing samples over time, being able to look at this process longitudinally from first to third grade. In the study, the authors concluded that "the best way to understand emerging bilingual children's biliterate development is to examine what they can do in both languages, as when looking at each language independently, one risks underestimating their full capacities" (p. 164). The authors argue that it is difficult to look at students' writing samples in Spanish and English isolatedly, as they do not represent all their weakness and strengths.

The longitudinal study by Witkowska et al. (2022) investigated syntactic complexity in monolingual children and children with English as an Additional Language (EAL), and also the relationship between vocabulary and syntax, using Mean Length of Utterance (MLU), Clausal Density (CD) and Complex Syntax type as methods of measuring the syntactic complexity. They found out that EAL children presented syntactically complex and diverse narratives which are equivalent to the monolingual participants. They also found that children with average vocabulary progressed at parallel rates. When analyzing children with low vocabulary scores, the results showed that EAL children developed at a faster rate than their monolingual peers, while the opposite was true for the high-vocabulary group. The researchers assume that "children with EAL might use language differently than their monolingual peers to achieve syntactic complexity and diversity" (p. 3), with skills comparable to their monolingual peers at early primary school. They

also indicate that storytelling is a useful tool for assessing children's knowledge of syntactic constructions.

Casanave (1994) found that L2's language development was associated with the production of longer and more complex syntactic clauses, measured by T-Units. In his study with intermediate Japanese students learning English, the author aimed at finding concrete ways to measure students' writing development. T-Unit analyses demonstrated that the writing of all the students changed over time, in different ways. The author highlights, however, that the improvement in students' writings should not be measured only quantitatively.

Following the same lines, Ortega (2003), based on a review of 25 studies involving L2 writing production, found that mean length of a sentence, mean length of T-Unit, mean length of clause and clauses per T-Unit were reliable indicators of L2 proficiency levels for adult writers. The author concluded that the relationship between L2 proficiency and L2 writing syntactic complexity varied across studies, depending on the context of investigation: second or foreign language.

In the studies reported in the present dissertation, the SI (number of clauses divided by the number of T-Units) will be used as a measure of syntactic complexity. One of the advantages of using this measure is that it does not matter if the student's production is short or long, since it assesses how well-connected the text is regardless of the size of the text. As younger students normally produce shorter texts, compared to older students, SI seems to be a good measure to prevent this difference in text size from impacting the results of the studies. In addition, SI is a simple measure and, for the analysis to be conducted, one needs only to know what a clause is. Furthermore, it can be used in different languages, which is the case of our study, in which oral and written texts in Portuguese and in English will be assessed through the counting of T-Units and their subordination index.

Ever since Hunt introduced the measure of T-Unit, many studies have adopted it as a method of analysis. Researchers have also started to use different measures related to or derived from T-Units, such as Mylläri (2020), who adopted seven syntactic complexity measures, out of which three were based on T-Unit count – mean length of T-Unit, mean number of T-units per sentence, and mean number of clauses per T-Unit. The author evaluated words, clauses, sentences, and T-units as production units in written learner language using a corpus of 352 L2 Finnish texts (28,813 words). Myllari (2020, p. 1) aimed at exemplifying how these categories are difficult to be determined and how hard it is to fit written learner language into these classifications. For Myllari (2020), defining the production units used and making visible the exceptions allowed, or the amount of data omitted, is extremely important when reporting ones' findings.

O'Donnell (1967) used T-Unit length as one of the measures while studying the development of writing and speech in schoolchildren and found that "several of the development trends observed for writing also hold true for speech - and hold even in the earliest grades." (O'Donnell, apud Hunt, 1970, p. 8). Also, he reported that the number of subordinate clauses per T-Unit increased at every grade interval, which is something Hunt (1970, p. 9) also pointed out.

> There is evidence to believe that throughout the school years, from kindergarten to graduation, children learn to use a larger and larger number of sentence-combining transformations per main clause in their writing. (...) In schoolchildren's speech the same tendency appears to exist up to the seventh grade, and future investigators may find that the tendency continues through the later grade.

Data from Hunt (1965), Klecan-Aker and Hedrick (1985), and Loban (1976) also showed an increase in the Subordination Index with age up to 12th grade both in conversation and narrative tasks, reinforcing the measure as an indication of linguistic development.

In a recent study, Frizelle et al. (2018) analyzed the narratives produced by 354 students, from 4 years to adulthood, showing that the number of clauses per utterance increased steadily. Their study also compared two narratives produced by the same participant: a narrative based on a sequence of images and a narrative recalled from memory, without the support of the images. The authors' findings showed that narratives generated by the task based on the sequence of images may underestimate syntactic competence in children under 5.

As it has been shown, syntactic complexity is a measure that has been used for a long time, proving to be adequate for its purposes. Furthermore, the choice of the task is another point to be considered, since narratives are usually used with children and are more likely to elicit complex language and full sentences, as compared to conversational samplings, for instance (Southwood; Russell, 2004; Gillon; Miller, 2004). In the studies reported in the present dissertation, SI measures will be correlated with other linguistic and cognitive measures and also with graph attributes through the analyses of oral and written narrative samples produced by bilingual children on the basis of sequence of images used to prompt their productions.

2.5 CONNECTING THE DOTS

What is the relation between what has been discussed so far? How can these different areas work together to reach our goal? This is what is going to be discussed in the next chapters. Starting from the idea that learning to read and write in two languages simultaneously does not present a problem to the students, and having the purpose to show it with Brazilian data and reality, this research was divided into three main studies in order to support our hypotheses.

Based on the literature review, we came across the most recognized studies on bilingualism and biliteracy, developed mainly in the northern hemisphere, representing a reality diverse from ours. The outcomes were mostly positive regarding bilingual education, but again, are not representative of our situation. A lack of studies with school children was also noticed, opening a field of needed research.

With that in mind, two questions needed to be answered so as to have a picture of the researches conducted with bilingual school children in the last decade, being: "What studies on the development of writing in bilingual children undergoing biliteracy are being carried out?" and "What are the conclusions of these studies regarding the development of writing skills in two languages?" The answer to these two questions will be presented in the next chapter, which brings studies on biliteracy with children. As will be shown in the chapter, there is a lack of research on biliteracy with school-aged children. Also, studies that encompass different realities and languages are another point made clear during our search for studies in this area.

With the need of specialized research in the area, a search for modern, low cost and non-invasive methodologies took place. We knew that T-units had been largely used in studies regarding syntactic complexity and that it could be an important analysis in our study. At the same time, we came across recent studies in language development using a new software - SpeechGraphs - which was able to quantify speech disorganization (Mota et al., 2012; 2014), but also served as a good predictor for measures of memory, analyzing whether such a relationship could be correlated with reading (Mota et al., 2019).

With the idea of having a study in a natural environment and with children that actually represent our reality in terms of language and kind of bilingual education, we developed our pilot study, in order to study biliteracy in school children and to analyze the effects of this process in both languages, Portuguese and English. This study will be presented in chapter 4. The pilot study was developed during the pandemic, 2020, during online classes.

Results from the pilot study were promising and led us to continue the research in the following year (2021), this time with classes being held at school and with a larger number of students and more grades involved. The research followed the same processes of the pilot study, but with some advances. Fortunately, students were back to school and all the data were collected at school, with three different age-groups: fourth, fifth and sixth graders. Another contribution of this study is the analysis of oral texts, in both languages. So, besides collecting and analyzing the students' productions in written texts, an analysis of their oral performance was added. Two groups of students were accompanied during the years of 2020 and 2021, in a longitudinal perspective.

In the next chapters, we present the studies mentioned above, that were designed, applied and completed during the last 2 years. The first brings a literature review of research in the area, presenting studies on biliteracy with children. The second brings the data and analysis made during the pilot project for the thesis, with data collected on-line in 2020, with about 50 children. The third presents a study conducted with 118 students, based on the pilot study, not only on written productions, but oral as well.

4. STUDY 2 — THE EFFECTS OF EARLY BILITERACY ON THOUGHT ORGANIZATION AND SYNTACTIC COMPLEXITY IN WRITTEN PRODUCTION BY 11-YEAR-OLD CHILDREN

4.1 INTRODUCTION

Estimates suggest that the majority of the world's population is bilingual (García & Cepeda, 2016; Kroll & Dussias, 2017). In Brazil, even though Portuguese is the only official language, the country possesses enormous language diversity, with around 330 languages being used daily, among which 274 indigenous languages (according to IBGE Census, 2012) and 56 immigration languages (Altenhöfen, 2013), in addition to Brazilian Sign Language (LIBRAS), recognised as one of the Brazilian languages since 2002.

In recent years, the number of bilingual schools has significantly increased in the country, most of them offering content classes being taught in English in addition to the regular curricula developed in Portuguese. There is, however, a critical lack of legislation regulating such an offer and a relevant scarcity of studies that investigate language processing and teaching methods considering the Brazilian context, in which children are first exposed to a prestigious language at the moment they enter school and need to develop reading and writing skills in a language they are not familiar with.

Within this context, it becomes imperative to better understand bilingual children's reading and writing development in order to design instructional pedagogies that contribute to supporting their growth as readers and writers considering the Brazilian bilingual schools' context. To fill out this empirical gap, the present study aimed at investigating the effects of biliteracy on the levels of thought organization (connectivity measures) and syntactic complexity in the written production in Portuguese and English in a group of children enrolled at a bilingual school located in the south of Brazil.

4.2 BILITERACY

In many bilingual schools in Brazil, the situation is similar to the one described by Petitto et al. (2013), since children enter school and may find themselves in a situation where they struggle to learn how to read and write in a language that they do not speak or use. In the case of these Brazilian kids, it is not known if the results regarding their abilities in reading and writing in both languages are as good as those of children who are only exposed to Portuguese, which fosters doubts as to whether schools should have children learn to read and write in their L1 first and

subsequently in their L2, or if it is not detrimental to have them develop reading and writing in both languages at the same time.

Regarding the discussion related to whether children should learn to read and write in two languages simultaneously or in sequence, the evidence in the previous literature is mixed. In one vision of sequential literacy, it is postulated that the development of literacy in the second language should not start before the child has developed the ability to speak, read and write in the L1 (Wong et al., 1986). On the other hand, there are some researchers who argue in favor of schools promoting the simultaneous development of reading and writing skills in the two languages of the children, even in situations in which children have not yet fully developed their L2 oral skills (García, 2006). This position is also defended by authors such as Edelsky (1986), Hudelson (1984) and Dworin (2003), among others. In his work, Dworin (2003) heightens the bidirectionality of the development of literacy, in which children's transactions with two written languages helps mediate their learning in both languages in a flexible and dynamic way, meaning that what is held in one language impacts the other, according to the author. From this perspective, what is learned in L1 impacts L2 and what is learned in L2 impacts L1. A few studies can be taken as evidence of the bidirectionality of literacy development in bilingual children. The longitudinal study developed by Soltero-González et al. (2016), for example, compared two models of instruction for emerging bilingual students in the United States, one promoting simultaneous literacy (paired-literacy) and the other fostering sequential literacy. The simultaneous literacy model consisted of providing instruction in English and Spanish from the beginning of schooling, based on the assumption that the individual's languages develop together. The main goal of this approach is to promote bilingualism and biliteracy, avoiding the transition of students from the instruction given in Spanish to instruction only in English. In the United States, however, the most common approach in dual-language programs is sequential, not simultaneous. For instance, in a study carried out with 358 Spanish/English speakers, from the third grade of 13 schools in Salem, Oregon, Francis et al. (2006) show that the group that followed the simultaneous program had better results in writing and reading, corroborating the idea that simultaneous and planned literacy leads to better development of both languages. In addition, it also confirmed that simultaneous instruction in English and Spanish did not inhibit the children's development of written English nor Spanish; on the contrary: it showed signs of strengthening English literacy while developing Spanish literacy. Dressler and Kamil (2006) also call attention to the evidence of cross-language transfer of reading comprehension skills in bilingual children across typologically distinct languages, throughout the time and bidirectionally, that is, from L1 to L2, as well as from L2 to L1. Ahmadi and Mohammadi (2019) researched biliteracy instruction in young learners, analyzing the effect of prior L2 literacy (English) on L1 (Persian) literacy. The results indicated that young

learners who gained functional literacy in L2 prior to L1 were more fluent and accurate L1 readers. Other studies by the same authors concluded that students (first graders) that were exposed to both L1 and L2 literacy outperformed the monolingual students. Also, research in the field of psycholinguistics has brought data that confirm that phonological and syntactic awareness knowledge may be transferred between their linguistic systems (Fu, 2003; Kabuto, 2011; Kuo & Anderson, 2007).

Cummins (1981, 2017), in his Interdependence Hypothesis, provides a framework for the transferring of knowledge and skills between the two of the speaker's languages. The author proposes a common underlying proficiency construct that supports the interaction between the bilingual languages, enhancing the development of literacy-related skills. According to him, transfer across languages is bidirectional, and instruction based on only one of the students' languages will minimize the contexts in which language development occurs.

Within the discussion regarding biliteracy development, other aspects may also have a noteworthy impact on the process, such as the linguistic distance between the languages that are being developed, whether they involve similar or distinct writing systems and their levels of linguistic transparency. These aspects may affect the way children process and develop language; therefore, they should be taken into consideration in the design of instruction methodologies. Languages that possess transparent orthographies, such as Portuguese, Finnish and Spanish, for example, have a direct one-to-one mapping between print and sound, whereas languages with more opaque orthography, such as English and French, have an irregular mapping between print and sound, which is also the case of logographic languages, such as Chinese. For instance, the study by Petitto et al. (2013), which was designed considering opacity between the children's languages, provided evidence that exposure to a language with a less opaque orthography, such as Spanish, helped children develop reading skills in English, a language that has a more opaque form of orthography. The authors came to the conclusion that instruction in both the children's languages during the same developmental period can bring reading advantages, not only reinforcing the importance of explicit exposure to both languages but also of providing phonological training in the two of the child's languages in the early school years.

Williams and Lowrance-Faulhaber (2018) analyzed 35 peer-reviewed studies on writing in bilingual children and showed that the development of literacy abilities in one language reinforces literacy knowledge and skills in the other of the children's language. The authors also emphasized that bilingual writing may progress in a similar way to what happens in the case of monolingual English speakers, even though bilinguals may follow distinct paths due to their unique language experience. According to the authors, the children were not confused; on the contrary, they used their linguistic knowledge bidirectionally with competence. These results

reinforce the idea that developing literacy-related skills in more than one language does not bring any harmful effects to children but, on the contrary, helps them develop reading and writing abilities more fully.

Many different studies have provided data showing that bilingual children often surpass monolingual children on tasks that assess literacy-related abilities when their both languages overlap in their writing system (Friedenberg, 1984; Da Fontoura & Seigel, 1995; Abu-Rabia & Siegel, 2002). Along the same lines, Bialystok et al. (2005) argue, for example, that the extent to which bilingualism affects literacy acquisition may depend on the similarity between the two language systems involved. But whether these results apply to contexts in which one of the children's language is not present in their everyday lives at home, that is, when they are learning a second language solely at school, is still an open question.

Along the same lines, Groff & Bellamy (2020) conducted a study with fourth-graders, speakers of P'urhepecha and Spanish. Not only did the results show that students produced better and longer texts in their L1, P'urhepecha, but they also demonstrated that children used richer vocabulary and a larger variety of verbal tenses. In addition, the study revealed that they found creative ways to represent both languages in both oral and written texts, reinforcing the advantage of providing opportunities for simultaneous development of both languages in educational contexts.

Finally, in the past few years, a large group of biliteracy researchers have been emphasizing the need for strategies that assess emerging bilingual children's biliteracy skills in a more holistic form, instead of adopting monolingual reading and writing assessments to design guidelines for literacy instruction. The Literacy Squared Biliterate Writing Assessment (Escamilla et al., 2014) was designed to fill out this gap and aims at capturing what bilingual students actually know in terms of literacy, assessing what they can do across and within languages in their two languages. According to this perspective, the language knowledge and abilities of bilingual children can never be appropriately measured or understood if these children are not assessed bilingually (Gort, 2006; Butvilofsky et al., 2017). Escamilla et al. (2014), for instance, present evidence that when students' languages are assessed from a holistic perspective, it is possible for teachers to better understand the way children operate with their two languages, contributing to leading children to greater outcomes in terms of biliteracy development.

In a very recent study, Butvilofsky et al. (2021) documented and analyzed the writing production of 29 second-grade bilingual students in Spanish and English. The authors qualitatively analyzed three sets of writing samples from students who had been identified as having poor or very poor reading scores in a traditional assessment. However, a holistic analysis of these students' productions revealed a better understanding of the complexity of their biliteracy development and

showed that the knowledge that students had in one language was often applied across languages or bidirectionally.

Taking into consideration the extensive research on biliteracy in other countries and the lack of research taking into consideration the Brazilian bilingual education context, we set out to investigate the development of writing skills in bilingual children from a specific context in Brazil, a bilingual school in which they go through the process of biliteracy, being Portuguese their L1, which is fully developed before they arrive at school, and English their emergent L2.

4.3 GRAPH ANALYSIS

Recently, network science and graph theory have gained increasing attention in the fields of neuroscience (Mota et al., 2012, 2014), psycholinguistics (Luz, 2018) and education (Mota et al., 2016, 2019). Mota et al. (2016, 2019) have conducted a series of experiments relying on graph analysis in order to explain cognitive development in healthy children as they progressed in their educational path through elementary school. Results showed that children whose oral memory reports generated graphs with more unique nodes (that denotes a larger vocabulary), more connections and fewer repetitions were the ones who scored higher in the cognitive and academic assessments, thus demonstrating the predictive power of graph analysis.

Graph attributes, particularly the Largest Connected Component (LCC) and the Largest Strongest Connected Component (LSCC) have been used to demonstrate patterns in oral and written texts which successfully distinguish between groups and evidence lack of connectedness as a strong indicator of cognitive development (Mota et al., 2016, 2019). According to Mota et al. (2016), the largest set of nodes directly or indirectly linked by some path is defined as the LCC, and the largest set of nodes directly or indirectly linked by reciprocal paths, in a way that all the nodes in the component are mutually reachable, is characterized as the LSC. In this sense, the LSC tends to be a stricter and more powerful predictor of connectivity since it closes a long-range word repetition cycle.

In the case of written language, one study of particular interest is the one developed by Luz (2018), who employed graph analysis to investigate patterns of connectedness in texts produced by good, bad, and dyslexic readers. The writing task required children to produce a story based on a comic strip without time or length constraints. The author compared graph parameters of texts to confirm that graph attributes were effective in sorting out good, bad and dyslexic readers, revealing patterns of textual connectedness, measured by a number of nodes and edges, LCC and text density.

It is relevant to note the innovative character of the present study, which is the first to adopt graph analysis in the investigation of writing development in bilingual children. It is argued here that such analysis, which makes use of low-cost, feasible and ecological assessment tools, may help provide important information regarding the development of second language oral and written production in bilingual children, which, in turn, can be used to design better intervention strategies in the near future.

4.4 METHOD

4.4.1 General Objective

The main goal of this study was to investigate the effects of bilingualism and biliteracy on the levels of thought organization (connectedness measures) and syntactic complexity in the written production in Portuguese and English in a group of fifty students, around 11 years old (M=10.7), enrolled in a bilingual school in the south of Brazil. This research proposal was approved by the Ethics Committee of the Federal University of Rio Grande do Sul under protocol number 5.290.938.

4.4.2 Specific Objectives

The specific objectives of the present study are as follows:
(1) To verify whether there is a difference in the L1 and L2 connectedness attributes (LCC and LSC), generated by the children's written production;
(2) To verify whether there is a difference in the L1 and L2 syntactic complexity measures (T-Units), generated by the children's written production;
(3) To explore the correlation between connectedness attributes and syntactic complexity measures in an attempt to verify whether graph analysis can serve as a potential tool to assess bilingual linguistic proficiency.

4.4.3 Hypotheses

In order to pursue the specific objectives, the following hypotheses were formulated: (1A) We expected to find a significant difference in the L1 and L2 connectedness attributes generated by the children's written production, with an advantage towards the participants' first language. In other words, connectivity measures were expected to be higher in Portuguese than in English texts.

(1B) Despite the L1 advantage in the written production, we expected to find a positive correlation between the connectivity attributes (LCC and LSC) in L1 and L2, signaling a parallel between thought organization expressed in the written production in the two languages.

(2A) We expected to find a difference in the L1 and L2 syntactic complexity measures (T-Units), generated by the children's written production, with an advantage towards the children's first language.

(2B) Despite the L1 advantage in the written production, we expected to find a positive correlation between measures of syntactic complexity (T-Units) in L1 and L2 texts, indicating that linguistic development occurs in a parallel fashion in the two languages of the bilingual children.

(3) We expected to find a positive correlation between the connectivity attributes (LSC) and syntactic complexity (T-units) in the L1 and L2 written productions, indicating that graph analysis can serve as a potential tool to assess bilingual linguistic proficiency.

4.4.4 Participants

A total of 50 typically developing children (19 male and 31 female, 10-11 yo in 2020, mean age 10.7) enrolled in 5th and 6th grades in a bilingual school located in the south of Brazil were invited to take part in the study. The children's home language is Portuguese, but they have been taught classes both in English and in Portuguese for at least five years and had already been screened for proficiency when they were in 4th grade (Cambridge Starters Exam). At the time the data was collected, the participants had had at least five years of bilingual education, with 10 hours of English per week (out of a total of 33 class hours), including English lessons and also lessons taught in English. English is not spoken in the community, even though it is present in social media, video games, songs, etc. In this case, it is a second language being added to the students' repertoire. The language students use to interact with each other is Portuguese. During English lessons, however, there is substantial use of the L2, which increases according to the students' progress.

4.4.5 Data Collection Procedures

Data was collected in August of 2020, during online classes, in two different moments, two weeks apart. There were three groups participating: one class of 6th graders and two of fifth graders. During their regular classes, each group was divided into two groups in alphabetical order. Participants were asked to create a narrative based on a sequence of five images (Cambridge Assessment, 2018), one in English and one in Portuguese, in a counterbalanced order. The first

group of students received the English version of the writing production, whereas the second group received the Portuguese version of the task. After two weeks, the same groups were kept but received the task in the other language, that is, the ones who had gotten the English version then got the Portuguese one, and vice-versa.

Students were asked to write a narrative (at least 200 words) based on a sequence of five images. The teacher (one of the researchers) started the class on the Google Meet platform by greeting the students. After that, she explained to students that they would not be evaluated but that they were expected to write the text according to the instructions given. The teacher reminded students that the story needed to be written according to the images and that they could create elements to enrich their stories. They needed to follow the sequence and facts of the images presented, though. As participants had to write at least 200 words, an explanation of how to use the word counting tool was given to them. Students opened their work in the Google Classroom and carried out their activity in silence, with the camera on. The teacher observed students during the task and answered questions. Examples of the picture sequences are shown in Figure 10 below.

Figure 10 - Cambridge Assessment, 2018

4.4.5.1 Data Analysis Procedures

The analysis of thought connectedness in both languages was performed considering the connectedness attributes generated by the computational tool Speech Graphs (Mota et al., 2014, 2016, 2019), and syntactic complexity was measured considering T-Units (Hunt, 1965). Both procedures are detailed below.

4.4.5.2 Graph Analysis

The students' original narratives were saved to a text file in which no changes were made. The free software Speech Graphs (Mota et al., 2014, available at http://neuro.ufrn.br/softwares/speechgraphs) was used to analyze the texts, representing each text in a distinct graph. A speech graph is a graphic representation of the sequential relationship of words in a verbal text, in which each word constitutes a node and the sequence between successive words constitutes what is called a direct edge. The software is also able to calculate several attributes per text file, including nodes (N) and edges (E), which represent the number of elements; repetitions of links between nodes and cycles of nodes that appear in the graphs, such as parallel edges (PE), repeated edges (RE), loops of one, two and three nodes (L1, L2 and L3); connectedness measures (LCC= largest connected component, LSC= largest strongly connected component) and others.

4.4.5.3 T-Units Analysis

Hunt (1965) first came up with the definition of a T-Unit, explaining it as the junction of a dominant clause and its dependent clauses. As the author points out, a T-Unit is the main clause and all subordinate clauses that are attached to it. T-Unit analyses have been mainly used in the areas of discourse, involving both written and spoken language, for example, in studies analyzing second language writing errors (Palmer, 2006; Ströbel et al., 2020; Myllari, 2020). In the present study, the T-Units analyses provided us with linguistic information and are related to the graph analysis data with the goal of potentializing the use of graphs to investigate the students' written production in both languages. These measures were compared to the graphs' characteristics in order to assess their correlations.

4.5 RESULTS

The results are presented below, taking into consideration our initial hypotheses stated in the Methods section. Table 3 displays the descriptive data which were analyzed in the comparison of the levels of syntactic complexity and thought organization in written production in Portuguese and in English by the children in our sample.

Table 3 - Descriptive statistics of measures of connectedness and syntactic complexity

Descriptive Statistics of Measures of Connectivity and Syntactic Complexity

		L1 Portuguese		L2 English	
	n	M	SD	M	SD
LSC	50	111.52	38.07849	84.42	25.91871
LCC	50	116.0	37.9121	88.23	25.59035
T-units	50	21.80	9.30657	18.6	6.298688

Note: n = number of participants; M = mean; SD = standard deviation; LCC= Largest Connected Component; LSC = Largest Strongest Connected Component; T-units = measure of syntactic complexity

In order to address our first hypothesis, Figure 11 displays the results from the Wilcoxon analysis comparing connectedness measures (LCC and LSC) in L1 and L2. As expected, there was a significant advantage for the texts written in Portuguese, the children's native language (LSC = W = 1193.5, *p*-value = 4.094e-06; and LCC= W =1231, *p*-value = 5.195e-06). In other words, connectedness measures were higher in Portuguese, as expected.

Figure 11 - Connectedness measures (LCC and LSC, respectively) in L1 (1) and L2 (2)
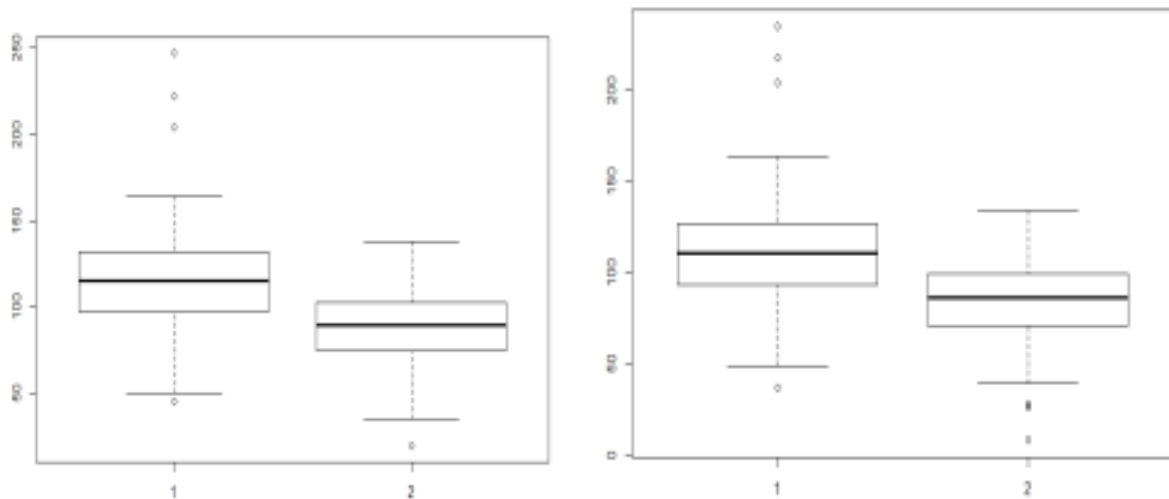
Our next step was to explore the correlation between connectedness measures (LCC and LSC) in L1 and L2. As displayed in Figure 12, a significant positive correlation between connectedness measures (LCC= rho 0.5171688, *p*-value = 0.000102, and LSC= rho 0.5511402, *p*-value = 2.775e-05) in L1 and L2 was found, indicating that linguistic development seems to occur in parallel and in the same proportion in the two languages of the bilingual child, as we expected.

Figure 12 - Correlation between connectedness measures (LCC and LSC) in L1 and L2
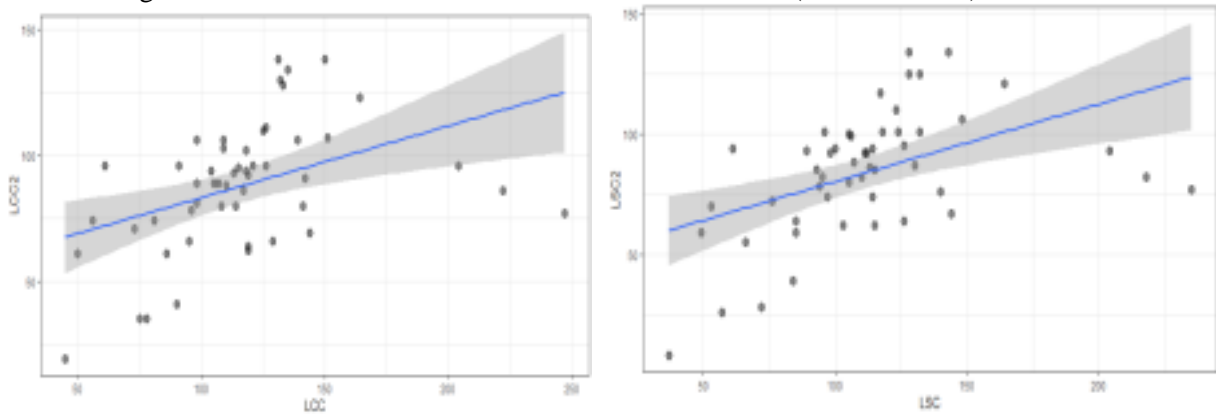
In order to illustrate the correlations that we have just reported, we bring the two graphs in Figure 13, which respectively show the high connectedness scores (LSC) of a participants' written text in L1 and in L2.

Figure 13 - Connectedness scores (LSC) of high-connected texts in L1 and L2 of the same participant

L1 = 150   *L2 = 138*
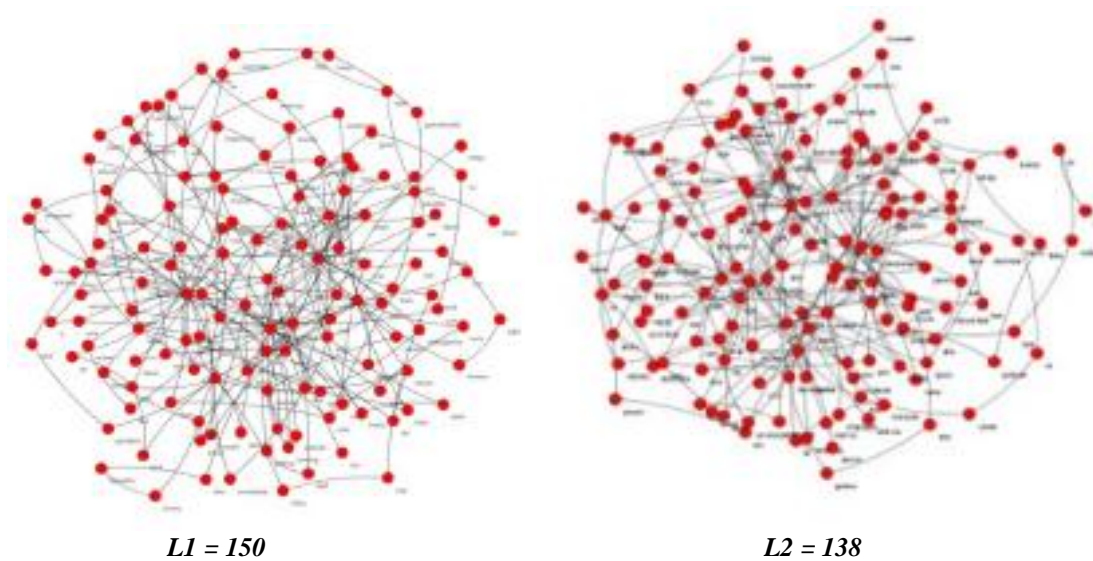
Conversely, the graphs shown in Figure 14 represent the results of a participant whose connectedness levels (LSC) were low both in Portuguese and in English.

Figure 14 - Connectedness scores (LSC) of low-connected texts in L1 and L2 of the same participant

L1 = 90   **L2 = 41**
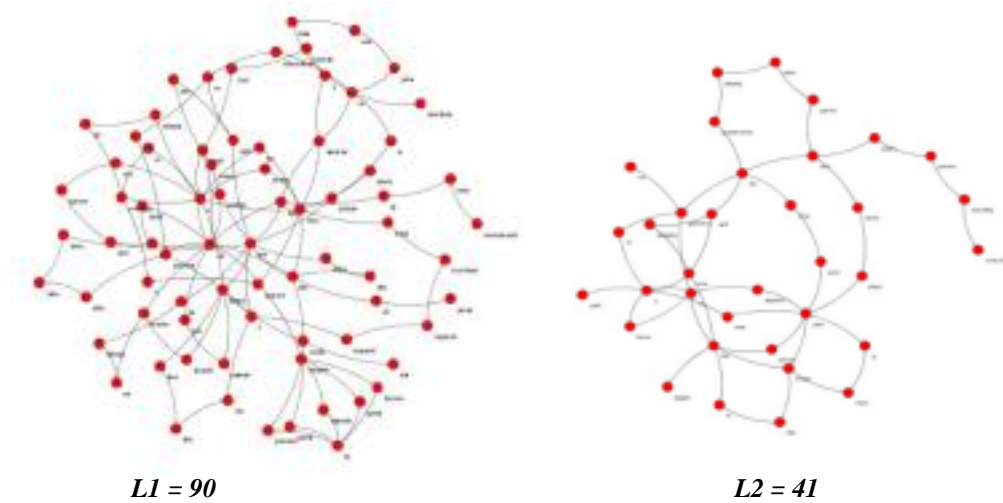
The graphs in Figures 12 show the analysis of a participant with high connectedness in L1 who also displayed high connectedness in L2. On the other hand, a student with low connectedness in L1 also showed low connectedness in L2. These results are in line with the findings of Abu-Rabia and Siegel (2002), reporting that children with poor reading skills in their L1 will exhibit weakness in their L2, and of Durgunoglu (2002), claiming that children with high performance in L1 will attain better results in L2.

An analysis of the participants' graphs indicates that the students' connectedness measures occur in the same proportion in both languages, with an advantage for their L1. These results are similar to the ones shown in previous investigations that demonstrate that L1 and L2 seem to develop together, in parallel, without hindering any of the languages and transferring linguistic skills from one language to the other bidirectionally (Dworin, 2003; Cummins, 2017).

Our second set of hypotheses (2A and 2B) dealt with the comparison between the means of syntactic complexity measures, and differences between the textual productions in the two languages were expected, with a significant advantage for the texts written in Portuguese children's native language. In other words, it was expected that Wilcoxon test results comparing the two means returned a significant difference, with a greater number of T-Units in the Portuguese texts, compared to the English productions, since it is the children's home and community language. This hypothesis was also confirmed, with a greater number of T-Units in Portuguese (T-Units = W = 738.5, p=0.03266), as Figure 15 shows. Again, the fact that their scores in L1 were higher than in their L2 does not alarm us, as their overall exposure to English (L2) is much lower than to Portuguese (L1).

Figure 15 - Syntactic complexity measures in L1(1) and L2(2)
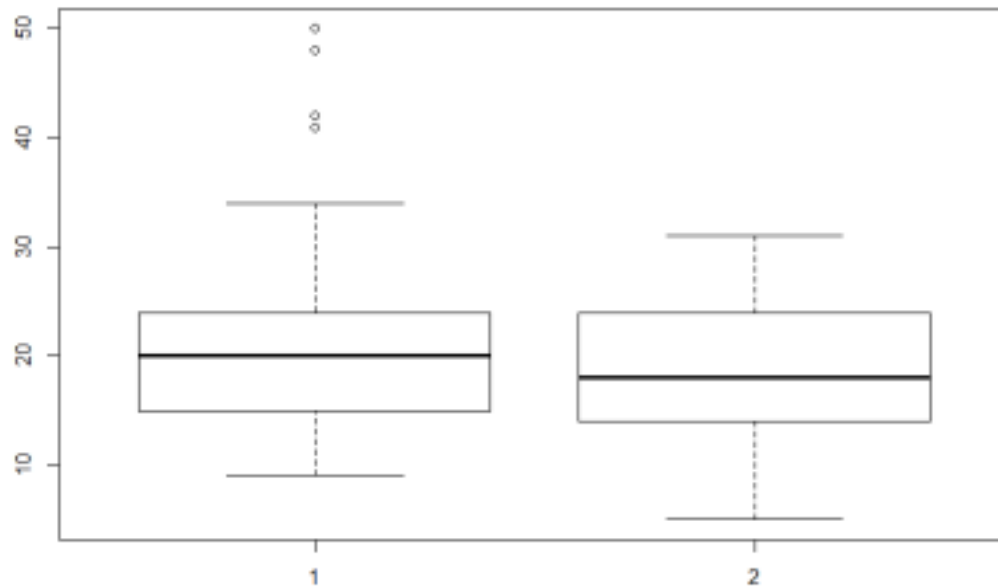
Our final hypothesis predicted a significant positive correlation between the measures of connectedness (LSC) and syntactic complexity (T-Units) in the participants' production in both L1 and L2. The positive correlations found between graph connectedness and complexity measures in L1 (S=4657.4, rho 0.7763569, *p*-value=3.479e-11) and L2 (S=10278, rho 0.5064645, *p* value=0.0001749), displayed in Figure 16, indicate that, in the present study, graph analysis served as a potential tool to assess the level of syntactic complexity of children in L1 and L2.

Figure 16 - Correlation between connectedness (LSC) and syntactic complexity (T-units) measures in L1 and L2
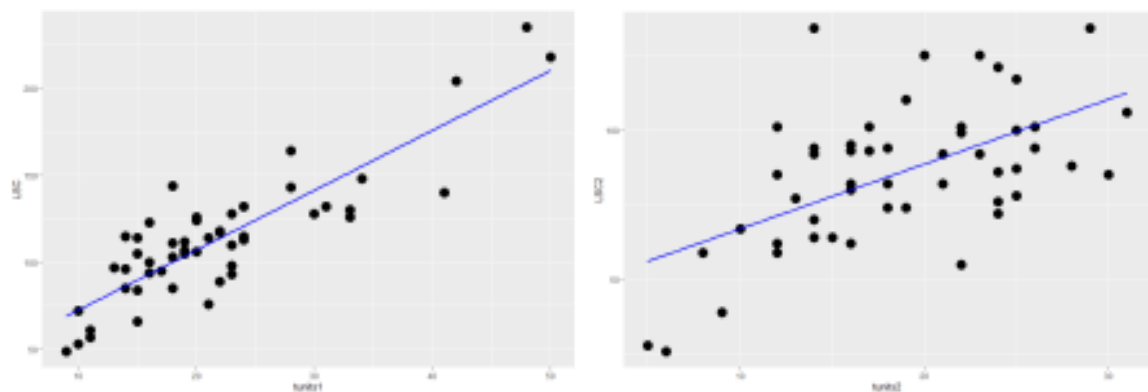
4.6 DISCUSSION

The results from both analyses – graph analyses and T-Units –revealed an advantage for the L1 Portuguese written production, as expected, with children obtaining higher connectedness measures and a larger count of T-Units in their mother tongue. Our first hypothesis was confirmed,

with higher connectedness measures in the Portuguese texts, which we interpret in terms of the children's greater exposure to Portuguese compared to English. It is the language they use to communicate daily, at home, with their friends and family. Even though there was a difference between the two languages, we do not interpret it as a negative impact of their L2 on their L1. In addition, as expected, we found a positive correlation between the connectedness attributes (LCC and LSC) in L1 and L2, signaling a parallel between thought organization expressed in the written production in the two languages.

Regarding our second hypothesis, related to the syntactic complexity (T-Units measures), our findings indicate that linguistic development appears to occur in parallel and in the same proportion in the two languages of the bilingual group. This could be taken as an indication that, as proposed by Bialystok et al. (2005), bilinguals may transfer writing skills acquired in one language to writing production in the other. This is in line with other authors who suggest that skills related to literacy development possibly transfer across languages as bilingual children progress through the school years (Cummins, 2017).

The third hypothesis was also confirmed since we found a positive correlation between connectedness measures and syntactic complexity measures. These results are particularly relevant in the growing context of bilingual education curricula and programs, which has brought concern related to finding more efficient ways to assess the development of bilingual children's reading and writing skills and the effect language instruction and exposure have on such complex development. Therefore, in this paper, we argue for the adoption of naturalistic, low-cost and large-scale measures, such as graph analysis, which may have a particular value in assessing oral and written production in young bilingual children. The design of more appropriate instructional practices can immensely benefit from this kind of evidence.

These results confirm the more consistent development in the participants' dominant language, which is the one they use at home and in the community and also the one in which they have received most instruction. Interestingly, despite the predicted L1 advantage in written production, a direct relationship between thought connectedness measures and syntactic complexity in both languages was found. These results are interpreted as evidence that, as children advance in the development of more complex writing strategies in Portuguese, they seem to progress in their written production in English to the same extent. In addition, our results reinforce the importance of teachers assessing students' written production in their two languages considering their bilingual experience, since the two languages of the bilingual are constantly active and in competition (Kroll & Bialystok, 2013).

4.7 CONCLUSIONS

The current study aimed at investigating the effects of bilingualism and biliteracy on the levels of thought organization (connectedness measures) and syntactic complexity in the written production in Portuguese and English in a sample of 11-year-old children enrolled in a bilingual school in the south of Brazil. Following Escamilla et al. (2014), we argue that only by assessing children's reading and writing skills in both languages and considering their knowledge and use of two languages we are able to reach a better understanding of their real development toward biliteracy. Besides, analyzing children's trajectories across languages allows us to better help teachers to design teaching pedagogies that foster the development of students' full biliteracy.

Also, what we believe to be a great contribution to the area is the use of Speech Graphs as a tool to naturalistically assess cognitive and linguistic performance in the case of kids being taught in two languages at the same time. The opportunity to get to know so much about the way students write and to be able to analyze their writing in order to establish new goals for improvement seems to be fascinating. Besides, understanding how a fast, simple, and noninvasive evaluation of speech connectedness relates to cognitive and academic performance may contribute to developing better intervention strategies in naturalistic settings. In addition, there is a clear gap in the literature, which points to the need to continue investigating biliteracy development and indicates that perhaps conducting longitudinal studies can be more informative.

We also know that this study needs to be expanded so that we can compare, for example, the students' writings throughout different school years. By doing so, we can analyze the way their writing progresses in both languages. One of the limitations of the study is the fact that data were collected during online classes due to the pandemics, which made it impossible to guarantee that participants were fully engaged in the tasks. Also, we consider it is relevant to replicate the study in different schools, with students from other contexts of language instruction, in order to analyze the possible effects of distinct kinds of intervention strategies and different amounts of L2 exposure in writing performance.

Our next steps include analyzing not only the students' written production but also their oral texts in order to compare their levels of thought organization and syntactic complexity in oral and written language. Therefore, Chapter 5 presents the follow-up study, which involved a larger number of participants and was conducted after a few adjustments were made to the protocol used in the pilot study.

6 FINAL CONSIDERATIONS, LIMITATIONS AND PEDAGOGICAL IMPLICATIONS

The main objective of this dissertation was to investigate the effects of schooling in two languages, English (L2) and Portuguese (L1), on the written and oral productions of students from the fourth to the sixth grade of elementary school, cross-sectionally and longitudinally. To that end, we analyzed the effects of school bilingualism and biliteracy on the levels of thought organization (connectedness measures: long-range and short-range recurrence) and syntactic complexity in students' written and oral narratives, in Portuguese and in English. The evaluation of measures of connectedness attributes was done through graph analysis, with the computational tool *SpeechGraphs* (Mota et al., 2014, 2016, 2019), and the syntactic complexity's analysis was based on T-Units (Hunt, 1965) and Subordination Index (SI).

To accomplish this dissertation's main goal, we conducted two empirical studies involving writing and oral development in two languages (Chapters 4 and 5), and a systematic review on the topic (Chapter 3). The systematic review came to meet a literature demand, due to the limited number of studies focusing on the investigation of biliteracy development in school contexts that were found. In that context, it became important to carry out a very careful and rigorous literature search, to make sure it would be comprehensive.

The systematic review aimed at answering two guiding questions: "What studies are being carried out on the development of writing in bilingual children undergoing biliteracy?" and "What are the conclusions of these studies regarding the development of writing skills in two languages?". It was based on the search for articles in English, Spanish and Portuguese in two large databases: Institute of Education Sciences (ERIC) and Scientific Electronic Library Online (SCIELO), considering articles published between 2012-2021.

Considering the breadth of the search, in terms of searched words, year (2012-2021) and languages of publication (English, Spanish and Portuguese), we found a small number of studies concerning the development of writing in bilingual school children, with 11 articles matching our search criteria. There was a larger number (395) of studies involving writing, mostly involving adult participants, as well as children with learning difficulties or special needs, not focusing on children in the literacy phase or in Elementary School. Therefore, we found an important gap in studies carried out with typical and bilingual children.

In 5 out of the 11 studies mentioned above, the participants were English and Spanish speakers from schools in the United States linked to the "Literacy Squared" project, a biliteracy program designed to respond to the need to cultivate new theories about the development of literacy in two different areas. All these studies argue for the development of literacy skills in two languages, and support the analysis of students' writing abilities from a bilingual perspective,

focusing on what students are capable of doing rather than focusing on the skills they have not yet been able to develop. The other 6 studies, also with emerging bilinguals but involving other languages, also brought favorable results to biliteracy, reinforcing the idea that there is an interrelationship between both languages of the bilingual child. In all cases, the authors concluded that literacy practices benefit from the development of two languages and do not cause any harm to student learning. Finally, the analysis conducted showed overall favorable results to biliteracy, reinforcing the idea that there is an intrinsic relationship between the two languages of a bilingual child and that literacy practices are potentialized and foster academic language development.

Following the systematic review, Chapter 4 presents Study 1, the pilot study that was conducted, which aimed at investigating the effects of bilingualism and biliteracy on the levels of thought organization (connectedness measures) and syntactic complexity in the written production in Portuguese and English in a group of fifty students, around 11 years old (M=10.7), enrolled in a bilingual school in the south of Brazil. The participants' task was to write a narrative (containing at least 200 words) based on a sequence of five images, one in English and one in Portuguese, in a counterbalanced order. The pandemics brought many research challenges and data collection within contexts was one of them. Regardless, we managed to collect the narrative samples in both Portuguese and English in August of 2020, during online classes, in two different moments, two weeks apart.

The results revealed an advantage for the L1 Portuguese written production, as expected, with children obtaining higher connectedness measures and a larger count of T-Units in their mother tongue, which is also the main language of schooling. In addition, we also found a positive correlation between the connectedness attributes that were measured (LCC and LSC) in the two languages (L1 and L2), signaling an association between the levels of thought organization expressed in the participants' written production in the two languages. Regarding syntactic complexity (T-Units measures), our findings also indicated that linguistic development appears to occur in parallel and in the same proportion in the two languages of the bilingual children. These results are taken as an indication that, as proposed by Bialystok et al. (2005), Escamilla et al. (2014), Hopewell & Escamilla (2014), Hopewell & Butvilofsky (2016), and Butvilofsky (2017, 2021), bilinguals may transfer writing skills acquired in one language to writing production in the other.

In addition, the results from Study 1 also showed a positive correlation between connectedness measures and syntactic complexity measures, which has led us to argue for the adoption of naturalistic, low-cost and large-scale measures, such as graph analysis, as tools that may have a particular value in assessing oral and written production in young bilingual children. It is of relevance to note that our results confirmed the more consistent development in the

participants' dominant language, which is the one they use at home and in the community and also the one in which they have received most instruction. Interestingly, despite the predicted L1 advantage in written production, a direct relationship between thought connectedness measures and syntactic complexity in both languages was found. These results are interpreted as evidence that, as children advance in the development of more complex writing strategies in Portuguese, they seem to progress in their written production in English to the same extent.

Chapter 5 presents the results of Study 2, which was conducted in 2021, as a follow up to our pilot study, and also included the assessment of oral development in two languages. The study had the goal of investigating cross-sectionally and longitudinally the effects of bilingualism and biliteracy on the levels of thought organization (connectedness measures: long-range recurrence and short-range recurrence) and syntactic complexity on written and oral narratives in two languages (L1 = Portuguese, L2 = English) by children in the 4th, 5th and 6th grades of Elementary School immersed in a bilingual schooling context. This study involved a larger sample of participants (118 children) and data collection was conducted in person, since students had already returned to school after the worst phase of the pandemics. Because children were having regular classes at school, in addition to the oral narratives in Portuguese and English that were collected individually, it was also possible to include in the protocol a cognitive measure to screen participants (Raven's Colored Progressive Matrices) as well as a L2 proficiency measure.

By measuring students' productions in terms of thought organization (connectedness measures: long-range and short-range recurrence) and syntactic complexity (SI scores), we were able to have a picture of their development from the fourth to the sixth grades of Elementary school. Similarly to what was found in the pilot study, overall results indicated that, as children advance in the development of more complex writing strategies in Portuguese, they seem to progress in their writing production in English to the same extent, but with an expected advantage for their L1. These results are consistent with Montanari et al. (2016), who showed that "it took students an additional year to reach the level of competence in the minority language that was comparable to the one they had in English, the majority and also dominant language for most students" (Montanari et al., 2016, p. 54), since students' L1 productions showed higher LSC and SI scores, which was a result we had expected to find.

In the analysis of the students' L1 and L2 written and oral productions in 2021, a year after the first moment of data collection, we found that only the connectedness measures showed a positive correlation. The same analysis, but with the SI scores was not significant, which was something we did not predict. In this case, LSC data seemed to be more consistent, showing a positive correlation across the students' languages, which was higher in written than in oral productions.

Another major finding is the progression we can see from one school year to the other, in both languages, in the cross-sectional analysis, which shows the interrelated language-based processes that develop in parallel, as Gort (2019) and other researchers point out (Gort, 2012, 2019; Grosjean, 1982; Reyes, 2006). In our analysis, we have shown that as children progress, their connectedness measures and SI's scores go up in both languages in a somewhat parallel fashion, showing a growth from fourth to fifth and sixth grades.

As for their oral development, the results from the correlation analyses suggest that children advanced in both languages, but with a possible higher cognitive demand for L2 oral production. In general, oral texts seem to be easier for children in terms of cognitive demand, as they do not need to transcribe their words into written material. As Berninger and Swanson (1994) claim, the transcription process posits a higher demand in the individuals' WM capacity, which normally explains why oral stories told by children are more complete than written stories. This, however, was not what our data has shown. We believe that due to the fact that they are producing narratives in an L2, the language of the task presented a greater cognitive demand than writing itself. Another aspect to consider is the format in question: written versus oral. Traditionally, school evaluation is assessed through written texts, which might give students an advance in this matter.

Another aspect to be considered is that speech connectedness may be a marker of individual differences in the early stages of L2 production, when the ability to produce a well-connected narrative tends to depend on a lexical repertoire that is still in development. These findings are consistent with the pattern reported in the early stages of L1 literacy, where increased longer recurrences were also associated with literacy development (Mota et al., 2016; 2018).

As for the possible correlations of proficiency in L2 and connectedness and SI scores, results showed an intermediate correlation between connectedness and proficiency in L2, for oral and written productions. They are in line with Leandro's findings (2021), in which he argues that speech graphs can differentiate between two levels of L2 proficiency, considering it an alternative to the evaluation of L2 performance. Our results corroborated his and also amplified its scope, since we analyzed written narratives in L2 as well. Along the same lines, the analysis of proficiency and L2 syntactic measures also indicated a positive correlation, even though it was low.

When we looked at the longitudinal data, LSC and SI scores do not have similar results. SI results in 2020 Portuguese and English productions had similar outcomes, while in 2021, a greater difference was found, with Portuguese 2021 texts achieving a higher score when compared to Portuguese 2020, and also when compared to English 2021. In fact, English productions did not show a significant increase from 2020 to 2021.

Connectedness measures, on the other hand, showed a higher score for English productions, with L2 scores being higher in both years, contradicting our initial hypothesis. However, when comparing the growth from one year to the other (2020 to 2021), we saw that the growth is greater in L1. Advances in terms of writing production were more significant in their L1 whereas students' L2 writing production remained steady, not presenting significant growth in a year span. This data may suggest that advances in the L2 require more time to take place in relation to the L1. Research has also revealed that bilingual students' writing skills in each language develops at different rates (Gort, 2006). Also, it might indicate that the pandemic affected L2 development more than L1. One reason for that is the fact that online classes tend to be shorter than regular classes. Also, in general, students participation online was significantly lower than at school.

One of the next steps we intend to take is the analysis of each group, per school year, in a longitudinal way, in order to verify if there are changes in this behavior according to students' age and grade new data is being collected in 2022). The students who were analyzed in the longitudinal data are a subgroup of the participants. They are currently in the seventh grade and their writing tasks were collected when they were in fifth and sixth grades. We believe that pandemic years affected the younger groups more intensively in terms of writing development and these further analyses might help us verify that.

It is important to observe, however, that this study had several limitations regarding sample size and tasks, among other things, including COVID-19 pandemic. Our data collections started in 2020, when schools were closed in Brazil and in most countries as well. Without a perspective of returning to face to face classes, we decided to start the data collection online, during synchronous classes. Students were kind and cooperative and participated in most of the tasks, but data collection suffered the impact of missing students in some days, for example. This is, unfortunately, something we could not prevent from happening. Also, the impact the data collection format had on students' performance cannot be fully accountable for. Lack of concentration, fewer classes, slower development of abilities expected for each school year, fewer opportunities to use the L2 and also psychological effects caused by isolation are some of the problems students encountered at the time and that may have affected the results obtained here.

In 2021 most students were back to school but the impact suffered in 2020 probably reflected on 2021's student's cognitive, linguistic and psychological development. Also, health protocols were very strict and students would miss school if they had any cold symptoms. This is the main reason for some students not having completed all tasks.

There are other limitations that are not related to the pandemic, being proficiency one of them. In our study, we did not have a proficiency measure for Portuguese (L1). Besides, the

proficiency test for L2 may not have been as effective as we intended it to be, since it regarded only reading and writing activities and it was adapted to a kahoot game.

Another difficulty we faced was related to the different sizes of oral and written samples. In order to prevent this from happening, in future research, we find it relevant to establish a predetermined amount of production time for the oral task. In our research, we decided to let students speak as much as they wanted, while for the written task we allocated the time of one class for them to do each task. These procedures may have somehow affected the production of the narratives too, which has implications for generalizations of the obtained results.

Considering the writing tasks, this was one of the methodological decisions we made from the pilot study to this one, as we decided not to establish a minimum number of words. Students had the same time (one class, around 40-50 minutes) to write, but without a fixed number of 200 words (as they had in the pilot study). This decision was taken as to let the production as naturalistic as possible, since we noticed that in the pilot study students were too concerned with the number of words, using different strategies to reach this number.

Another limitation of this study is the fact that we tested children of medium and high socioeconomic level, which also limits the possibility of generalization. As it is known, socioeconomic level affects language development, by determining possibilities of schooling and access, especially in L2.

Furthermore, more studies need to be conducted using *Speechgraphs* in the analysis of written texts. The importance of looking at different data and analyzing the best methodological aspects of graph attributes for healthy and young populations is paramount. The study being conducted by Scholl (2022) may shed light to the discussion, as the author is looking at writing productions in L1 and L2 in two different text types.

Despite its limitations, we believe our study provides an important contribution to a better understanding of bilingual children's oral and written development in L1 and L2 in the context of prestige bilingual education in Brazil. Our next steps include collecting and analyzing students' performance in 2022, looking at the same participants' written and oral development throughout the period of three years. Also, an analysis of each group (4th grades, 5th graders and sixth graders) is planned to be done in order to capture the development of students in different stages cross-sectionally, but also in a longitudinal way. It would also be important to include a proficiency measure of students' L1. By doing that, it would be possible to analyze the data with a model of proficiency that accounted for both languages.

Another follow up study, already in motion, regards the comparison of students' development in different bilingual schools in Brazil, looking at their progress in both languages.

The data from Costa et al. (2022, submitted) might have an interaction with ours, expanding the analysis of 5th grade students being educated bilingually in two different states in Brazil.

One of the main contributions of this dissertation is directly connected to bilingual education. As a matter of fact, this study started in the classroom, with my students and I working together and it is of my greatest interest that it returns, somehow, to the classroom in order to benefit students and teachers. Firstly, it helps fight myths regarding bilingual education and biliteracy development as it shows that the development in one language does not harm the other. Secondly, it shows that languages develop together, both in written and in oral productions, which serves as an indicator that biliteracy is a possibility, doing so with data that represents bilingual education in Brazil. Thirdly, it calls teachers to look at this phenomena and urges them to analyze students' writing and oral development from a bilingual perspective, emphasizing the importance of a planned bilingual curriculum towards biliteracy.

The possibility of linking the innovation of computational methods, with a high degree of precision and rigor, to measures traditionally used in the studies of Psycholinguistics, opens new avenues of investigation and possibility of data analysis from more numerous and diversified samples. It is important to highlight the innovative character of the present study, which is the first to adopt graph analysis in the investigation of writing and oral development in bilingual children. The analysis makes use of a low-cost, feasible and ecological assessment tool to measure young bilingual children's growth as speakers and writers, which, in turn, can be used to design better intervention strategies in the near future.

Having in mind our general objective when we started this study and considering the scarcity of empirical research that portrays the development of writing in bilingual schooled children, specifically in prestigious bilingual schools, we consider that our initial goal was accomplished. Our research, which is representative of the Brazilian context, contributes to the development of the biliteracy process in prestige education schools, specifically involving two alphabetical languages: Portuguese and English, in a context where L2 is developed at school and L1 is the community and school language.

REFERENCES

Abraham, S. L. (2017). A teacher's inquiry into bringing in biliteracy in a fifth-grade English-only classroom. *Networks: An Online Journal for Teacher Research*, *19*(1), Article 2.

Abu-Rabia, S., & Siegel, L. S. (2002). Reading, syntactic, orthographic, and working memory skills of bilingual Arabic-English speaking Canadian children. *Journal of Psycholinguistic Research, 31*, 661-678. https://doi.org/10.1023/A:1021221206119

Loy, A., Steele, S., & Korobova, J. (2022). lmeresampler: Bootstrap Methods for Nested Linear Mixed-Effects Models. R package version 0.2.2. https://CRAN.R-project.org/package=lmeresampler

Ahmadi, N., & Mohammadi, N. D. (2019). On the sequence of biliteracy instruction in young learners: L2 then L1, or simultaneous L1 and L2? In Z. Flegar & I. Moritz (Eds.), *Children and languages today,* (pp. 83-98). Vernon Press.

Kassambara, A. (2021). Rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.0. https://CRAN.R-project.org/package=rstatix

Alexanderson, G. (2006). About the cover: Euler and Konigsberg's Bridges: A historical view. *Bulletin of the American Mathematical Society*, *43*(4), 567-573.

Alves, U .K., & Finger, I. (2023) *Alfabetização em contextos monolíngue e bilíngue*. Vozes Editora.

Altenhöfen, C. V. (2013). Bases para uma política linguística das línguas minoritárias no Brasil. In C. Nicolaides *et al*. *Política e Políticas Linguísticas* (pp. 93-116). Pontes Editores.

Anderson, T., & Boyer, M. (1970). *Bilingual schooling in the United States.* Southwest Educational Development Laboratory.

Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., & Duarte, J. L. M. (1999). *Matrizes Progressivas Coloridas de Raven: escala especial. Manual*. CETEPP.

Azevedo A. F., Buchweitz, A., Esper, N., Costa, A., & Franco, A. (2017) Neuroimagem funcional na dislexia do desenvolvimento em bilíngues de ortografias opacas e transparentes. In J. F. de Salles & A. L. Navas. (Org.). *Dislexias do desenvolvimento e adquiridas* (pp. 197-211). Pearson.

Bae, J. (2007). Development of English Skills Need Not Suffer as a Result of Immersion: Grades 1 and 2 Writing Assessment in a Korean/English Two-Way Immersion Program. *Language Learning: A Journal of Research in Language Studies*, *57*(2), 299-332. https://doi.org/10.1111/j.1467-9922.2007.00410.x

Baker, C. (2006). *Foundations of Bilingual Education and Bilingualism*. Multilingual Matters.

Baker, C., Wright, W. (2017). *Foundations of bilingual education and bilingualism* (6th ed.). Multilingual Matters.

Bauer, E., Gort, M. (Eds.) (2012). *Early biliteracy development*: Exploring young learners' use of their linguistic resources. Routledge.

Becker, A. (2006). *A review of writing model research based on cognitive processes*.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Lawrence Erlbaum Associates.

Bice, K., & Kroll, J. F. (2015). Native language changes during early stages of second language learning. *NeuroReport*, *26*, 966-971.

Bloomfield, L. (1933). *Language*. George Allen and Unwin.

Brasil. Lei nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. Diário Oficial [da República Federativa do Brasil], Brasília, 225, DF, *134*(248), 23 dez. 1996. Seção I, 27834-27841. https://www.jusbrasil.com.br/busca?q=Art.+62+da+Lei+de+Diretrizes+e+Bases+-+Lei+9394%2F96.

Brentano, L., & Finger, I. (2020). Biliteracia e educação bilíngue: Contribuições das Neurociências e da Psicolinguística para a compreensão do desenvolvimento da leitura e escrita em crianças bilíngues. *Revista Letrônica*, *13*(4). https://doi.org/10.15448/1984-4301.2020.4.37528

Bialystok, E., Mcbride-Chang, C., & Luk, G. (2005). Bilingualism, language proficiency, and learning to read in two writing systems. *Journal of Educational Psychology, 97*(4), 580-590. https://doi.org/10.1037/0022-0663.97.4.580

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, *10*, 210-223.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240-250. https://doi.org/10.1016/j.tics.2012.03.001

Butvilofsky, S.A.; Hopewell, S.; Escamilla, K.; Sparrow, W. (2017). Shifting deficit paradigms of latino emerging bilingual students' literacy achievement: documenting biliterate trajectories. *Journal of Latinos and Education*, *16*(2), 85-97. https://doi.org/10.1080/15348431.2016.1205987.

Butvilofsky, S. A., Escamilla, K., Gumina, D., & Silva Diaz, E. (2021). Beyond Monolingual Reading Assessments for Emerging Bilingual Learners: Expanding the Understanding of Biliteracy Assessment through Writing. *Reading Research Quarterly, 56(1)*, 53-70.

Cambridge English Language Assessment. *Cambridge English First for Schools*: Handbook for teachers. http://www.cambridgeenglish.org.

Casanave, C. (1994). Language development in students' journals. *Journal of Second Language Writing*, *3*, 179–201. http://dx.doi.org/10.1016/1060-3743(94)90016-7

Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*, *40*, 249-268.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. Feltovich, R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 121-130). Cambridge University Press.

Collins, B. A., O'Connor, E. E., Suarez-Orozco, C., Nieto-Castañon, A., Toppelberg, C.O. (2014). Dual language profiles of Latino children of immigrants: Stability and change over the early school years. *Applied Psycholinguistics*, *35*(3), 581-620. http://dx.doi.org.libaccess.sjlibrary.org/10.1017/S0142716412000513

Crossley, S. A., McNamara, D. S. (2014). Does writing development equal writing ability? A computational investigation of syntactic complexity in L2 learners. J*ournal of Second Language Writing*, *26*, 66-79.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*(2), 222-251.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students*: A theoretical framework (pp. 3–49). Evaluation, Dissemination and Assessment Center California State University.

Cummins, J. (2017). Teaching for transfer in multilingual school context. In O. García, A. M. Y. Lin, & S. May. (Eds.) *Bilingual and Multilingual Education* (3 ed.) (pp. 103-116), Cham.

De Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics, 13*, 1-24.

De Fontoura, H. A., & Siegel, L. S. (1995). Reading, syntactic and working memory skills of bilingual Portuguese English Canadian children. *Reading and Writing: An Interdisciplinary Journal, 7*(1), 139-153. https://doi.org/10.1007/BF01026951

de Groot, A. M. B. (2011). *Language and cognition in bilinguals and monolinguals*. Psychology Press.

De La Piedra, M. T., & Araujo, B. (2012). Transfronterizo literacies and content in a dual language classroom. *International Journal of Bilingual Education and Bilingualism*, *15*(6), 705-721.

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116*(15), 1-10.

Diamond, A. (2012). *Executive Functions*. 10.1146/annurev-psych-113011-143750

Diebold, A. R. (1964). Incipient bilingualism. *Language*, *XXXVII*, 97-112.

Dressler, C., & Kamil, M. L. (2006). First- and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners. Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 197-238). Mahwah: Lawrence Erlbaum Associates Publishers.

Durán, L. (2017). Understanding young children's everyday biliteracy: "Spontaneous" and "scientific" influences on learning. *Journal of Early Childhood Literacy*, *18*(1),71-96.

Durgunoglu, A. Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia, 52*, 189–204.

Dworin, J. E. (2003). Insights into biliteracy development: Toward a bidirectional theory of bilingual pedagogy. *Journal of Hispanic Higher Education*, *2*(2), 171-186. https://doi.org/10.1177/1538192702250621

Edelsky, C. (1986). *Writing in a bilingual program*: *Había una vez*. Ablex Publishing.

Elbers, L. (1982). Operating principles in repetitive babbling: A cognitive continuity approach. *Cognition*, *12*, 45-63.

Escamilla, K., Hopewell, S., Butvilofsky, S., Sparrow, W., Soltero-González, L., Ruiz-Figueroa, O., & Escamilla, M. (2014). *Biliteracy from the start: Literacy squared in action* (pp. 25-26). Caslon Publishing.

Ferguson, C. A., Houghton, C., & Wells, M. H. (1997). Bilingual Education: An International perspective. In B. Spolsky & R. Cooper (Eds.), *Frontiers in bilingual education* (pp. 159-174). Newbury House.

Field, A.; Miles, J.; Field, Z. (2012). *Discovering Statistics Using R*. SAGE Publications.

Field, A., Jeremy, M., & Zoë, F. (2012). *Discovering Statistics using R*. SAGE Publications Ltd.

Finger, I. (2015). Psicolinguística do bilinguismo. In L. S. Rebello & V. N. Flores (Orgs.) *Caminhos das letras: uma experiência de integração*. Ed. Instituto de Letras/UFRGS.

Finger, I., Brentano, L. S., Ruschel, D. (2019). E quando a alfabetização ocorre simultaneamente em duas línguas? Reflexões sobre o biletramento a partir da análise de textos de crianças bilíngues. *ReVEL*, *17*(33).

Finger, I., Ortiz-Preuss, E. (2018). A psicolinguística do bilinguismo: estudando o processamento linguístico e cognitivo bilíngue. In I. Finger & E. Ortiz-Preuss. *A dinâmica do processamento bilíngue*. Pontes.

Fishman, J. (1976). *Bilingual education: An international sociological perspective*. Newbury House Publishers, Inc.

Fishman, J. (1977). The social science perspective. In J. Fishman (Ed.), *Bilingual education: Current perspectives*, Social science (pp. 14-62). Center for Applied Linguistics.

Fishman, J. (2001). 300-plus years of heritage language education in the United States. In J. K. Peyton, D. A. Ranard, & S. McGinnis (Eds.) *Heritage languages in America: Preserving a national resource* (pp. 81-89). Center for Applied Linguistics & Delta Systems,.

Flores, N. (2016). A tale of two visions: Hegemonic whiteness and bilingual education. *Educational Policy*, *30*(1), 13–38.

Flores, N. (2017). Bilingual education. In O. García, N. Flores, & M. Spotti (Eds.), *Oxford Handbook of Language and Society* (pp. 525–44). New York: Oxford University Press.

Francis, D., Lesaux, N., & August, D. (2006). Language of instruction. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners; Report of the National Literacy Panel on Language Minority Children and Youth* (pp. 365-414). Erlbaum.

https://doi.org/10.4324/9781315094922

Friedenberg, J. E. (1984). The effects of simultaneous bilingual reading instruction on the development of English reading skills. *Acta Paedologica, 1*(2), 117-124.

Fu, D. (2003). *An Island of English: Teaching ESL in Chinatown*. Heinemann.

Gaies, S. (1980). T-Unit Analysis in Second Language Research: Applications, Problems and Limitations. *TESOL Quarterly*, *14*(1), 53-60.

García, O. (2009). *Bilingual Education in the 21st Century*: a global perspective. Blackwell Publishing.

García, O., Wei, L. (2014). *Translanguaging*. Language, bilingualism and education. MacMillan.

García, O., Lin, A. M. Y. (2017). Extending Understandings of Bilingual and Multilingual Education. In O. García, A. M. Y. Lin, & S. May. *Bilingual and multilingual education* (pp. 1-20). 3 ed. Springer International Publishing.

García, O. (2006). *Lost in transculturation: The case of bilingual education in New York City.* In M. Pütz, J. A. Fishman, & J. A. Neff-van Aertselaer (Eds.), *Along the Routes to Power: Explorations of the Empowerment through Language* (pp. 157-178). Mouton de Gruyter.

García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Wiley/Blackwell.

García, A. M., & Cepeda, S. S. (2016). *Mente bilingüe: Abordajes psicolinguísticos y cognitivistas*. Córdoba: Comunic-Arte.

Garcia, G. D. (2021). *Data Visualization and Analysis in Second Language Research*. Routledge. https://doi.org/10.4324/9781003032243

Gazzaniga, M., Ivry, R., & Mangun, G. (2009). *Cognitive Neuroscience: The biology of the mind* (3rd ed). Norton & Company.

Genesee, F., Lindholm-Leary, K. J., Saunders, W., & Christian, D. (2006). *Educating English Language Learners*. Cambridge University Press.

Genesee, F., Fortune. T. W. (2014). Bilingual education and at-risk students. *Journal of Immersion and Content-Based Language Education, 2*(2), 196-209. https://doi.org/10.1075/jicb.2.2.03gen

Gort, M. (2006). Strategic codeswitching, interliteracy, and other phenomena of emergent bilingual writing: Lessons from first grade dual language classrooms. *Journal of Early Childhood Literacy, 6*(3), 323-354. https://doi.org/10.1177/14687 98406 069796.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language, 27*(2), 210-223.

Gries, S. (2013). *Statistics for Linguistics with R: A Practical Introduction* (2nd Edition). De Gruyter Mouton. https://doi.org/10.1515/9783110307474.

Groff, C., & Bellamy, K. (2020). Biliteracy Development in Mexican Primary Education: Analyzing Written Expression in P'urhepecha and Spanish. *Language Learning Journal*, *48*(3),

285-299.

Grosjean, F. (1999). The bilingual's language modes. In J. L. Nicol (Ed.). *One Mind, Two Languages: Bilingual Language Processing*. Blackwell.

Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism, 2*, 5. Wiley-Blackwell.

Hamers, J., Blanc, M. H. A. (2000). *Bilinguality and Bilingualism* (2 ed). Cambridge University Press.

Hawkins, R. (2001). *Second language syntax: A generative introduction.* Wiley-Blackwell.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369–388. https://doi.org/10.1177/0741088312451260

Hélot, C. (2006). Bridging the gap between prestigious bilingualism and the bilingualism of minorities: Towards an integrated perspective of multilingualism in the French education context. In M. Ó. Laoire (Org.). *Multilingualism in educational settings* (pp. 49-72). Schneider Verlag Hohengehren.

Hopewell, S., Butvilofsky, S. (2016). Privileging bilingualism: using biliterate writing outcomes to understand emerging bilingual learners' literacy achievement. *Bilingual Research Journal, 39*(3-4), 324-338.

Hopewell, S., Escamilla, K. (2014). Biliteracy development in immersion contexts. *Journal of Immersion and Content-Based Language Education, 2*(2), 181-195.

Hornberger, N. H. (2003). Continua of biliteracy. *Bilingual education and bilingualism* (3-34).

Hudelson, S. (1984). Kan yu ret and rayt en Ingles: Children become literate in English as a second language. *TESOL Quarterly, 18*, 221–238.

Hunt, K. (1965). Grammatical structures written at three grade levels. *NCTE Research Report*, *3*. NCTE. https://files.eric.ed.gov/fulltext/ED113735.pdf.

Instituto Brasileiro de Geografia e Estatística (IBGE). (2012). *Censo Brasileiro de 2010*. IBGE.

IPOL. Lista de línguas cooficiais em municípios brasileiros. (2022). http://ipol.org.br/lista-de-linguas-cooficiais-em-municipios-brasileiros/

Jones, W. R., Stewart, W. A. C. (1951). Bilingualism and verbal intelligence. *British Journal of Statistical Psychology*, *4*, 3-8.

Kabuto, B. (2011) *Becoming biliterate: Identity, ideology, and learning to read and write in two languages*. Routledge.

Kassambara, A. *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. R package version 0.7.0. 2021. https://CRAN.R-project.org/package=rstatix

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy, & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp.

57–71). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203811122

Kirk, S. (2014). Addressing Spoken Fluency in the Classroom. In: T. Muller, J. Adamson, P. S. Brown, S. Herder. (Eds.) *Exploring EFL Fluency in Asia.* Palgrave Macmillan. https://doi.org/10.1057/9781137449405_7

Kroll, J. F., & Dussias, P. E. (2017). The Benefits of Multilingualism to the Personal and Professional Development of Residents of The US. *Foreign Language Annals, 50,* 248–259.

Kroll, J. F., & Dussias P. E. (2018). Language Use Across International Contexts: Shaping the Minds of L2 Speakers. *Annual Review of Applied Linguistics, 38*, 60–79. https://doi.org/10.1017/S0267190518000119

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497-514. 10.1080/20445911.2013.799170

Kuo, L. J., & Anderson, R. C. (2007). Conceptual and methodological issues in comparing metalinguistic awareness across languages. In K. Koda, & A. M. Zehler (Eds.), *Learning to Read Across Languages: Crosslinguistic Relationships in First-and-Second-Language Literacy Development* (pp. 39-67). Routledge. https://doi.org/10.4324/9780203935668-10

Leandro, D. C. (2020). *Planejamento pré-tarefa, memória operacional e produção oral em inglês como L2: um estudo exploratório usando análise de grafos* [Doctoral dissertation], Universidade Federal do Rio Grande do Norte.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on Fluency* (pp. 25–42), University of Michigan Press,.

Levelt, W. J. M. (1989). Speaking: From intention to articulation. Bradford/MIT Press.

Levelt, W. J. M. (1999). Producing spoken language: a blueprint of the speaker. In C. M., Brown, P. Hagoort. (Eds). *The Neurocognition of language*. Oxford University Press.

Levelt, W. J. M. (2019). How Speech Evolved: Some historical remarks. *Journal of Speech, Language, and Hearing Research*, *62*, 2926-2931.

Levshina, N. (2015). *How to do Linguistics with R: Data Exploration and Statistical Analysis*. John Benjamins. https://doi.org/10.1075/z.195.

Liberali, F., & Megale, A. H. (2016). Elite bilingual education in Brazil: an applied linguist 's perspective. *Colombian Applied Linguistics Journal*, *18*(2), 95-108.

Lindholm-Leary, K. J. (2014). Bilingual and biliteracy skills in young Spanish-speaking low-SES children: impact of instructional language and primary language proficiency. *International Journal of Bilingual Education and Bilingualism*, *17*(2), 144-159. 10.1080/13670050.2013.866625

Lindholm-Leary, K. J. (2016). Bilingualism and academic achievement in children in dual language programs. In E. Nicoladis, & S. Montanari. (Eds.). *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 203-223). APA Books.

Lindholm-Leary, K. J. (2018). Implications of research for dual language at the early childhood level. In M. B. Arias, & M. Fee (Eds.). *Profiles of Dual Language Education in the 21st Century*. Multilingual Matters.

Loban, W. (1976). *Language Development: Kindergarten through Grade Twelve*. NCTE Committee on Research, Report No. 18, 156 p.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474-496.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605-621. http://doi.org/10.1080/20445911.2013.795574

Luz, J. P. (2018). *Análise de grafos aplicada a produções textuais de alunos do ensino fundamental e seu potencial preditivo da dislexia do desenvolvimento* [Master's thesis], Pontifícia Universidade Católica do Rio Grande do Sul. http://tede2.pucrs.br/tede2/handle/tede/8158

Marcelino, M. (2009). Bilinguismo no Brasil: significado e expectativas. *Revista Intercâmbio*, *XIX*, 1-22. LAEL/PUC-SP.

May, S. (2017). Bilingual Education: What the Research Tell Us. In O. García, A. M. Y. Lin, & S. May. *Bilingual and multilingual education*. 3 ed. Springer International Publishing. p. 81-100.

Montanari, S., Simon-Cereijido, G., & Hartel, A. (2016). The development of writing skills in an Italian-English two-way immersion program: evidence from first through fifth grade. *International Multilingual Research Journal, 10*(1), 44-58.

Hervé, M. (2022). RVAideMemoire: Testing and Plotting Procedures for Biostatistics. R package version 0.9-81-2. https://CRAN.R-project.org/package=RVAideMemoire

Montrul, S. (2012). Is the heritage language like a second language? *Eurosla Yearbook*, *12*(1), 1-29. http://www.nhlrc.ucla.edu/media/files/eurosla-yearbook-2012-text-yz-rvu.pdf

Morais, J. (2013). *Criar leitores:* para professores e educadores. Manole. 154p.

Mota, N. B., Vasconcelos, N. A., Lemos, N., Pieretti, A. C., Kinouchi, O., Cecchi, G. A., Copelli, M., & Ribeiro, S. (2012). Speech graphs provide a quantitative measure of thought disorder in psychosis. *PloS One*, *7*(4). e34928. https://doi.org/10.1371/journal.pone.0034928

Mota, N. B., Furtado, R., Maia, P. P., Copelli, M., & Ribeiro, S. (2014). Graph analysis of dream reports is especially informative about psychosis. *Scientific Reports*, *4*, 3691. https://doi.org/10.1038/srep03691

Mota, N. B., *Weissheimer, J., Madruga, B., Adamy, N., Bunge, S. A., Copelli, M., & Ribeiro, S.* (2016)*.* A Naturalistic assessment of the organization of children's memories predicts cognitive functioning and reading ability. *Mind, Brain and Education*, *10*(3), 184-195.

Mota, N. B., Sigman, M., Cecchi, G., Copelli, M., & Ribeiro S. (2018). The maturation of speech structure in psychosis is resistant to formal education. *NPJ Schizophr, 4*(25). doi: 10.1038/s41537-018-0067-3.

Mota, N. B., Copelli, M., & Ribeiro, S. (2018). Graph theory applied to speech: insights on cognitive deficit diagnosis and dream research. In T. Poibeau, & A. Villavicencio. *Language, cognition, and computational models.* Cambridge University Press.

Mota, N. B., Callipo, R., Leite, L., Torres, A. R., Weissheimer, J., Bunge, S., Copelli, M., & Ribeiro, S. (2019). Verbal short-term memory underlies typical development of thought organization measured as speech connectedness. *Mind, Brain and Education*, *14*(1), 51-60.

Moura, S. A. (2010).Educação bilíngue e currículo: de uma coleção de conteúdos a uma integração de conhecimentos. In C. H. Rocha, J. R. A. Tonelli, & K. A.da Silva, (Eds.) *Novas perspectivas em linguística aplicada* (pp. 269-295). Pontes Editores.

Mylläri, T. (2020). Words, clauses, sentences, and T-units in learner language: Precise and objective units of measure? *Journal of the European Second Language Association, 4*(1), 13–23. https://doi.org/10.22599/jesla.63

O'Donnell, R. C., Griffin, W. J., & Norris, R. C. (1967). *Syntax of Kindergarten and Elementary School Children: a transformational analysis. NCTE Research*, *Report No. 8*. NCTE.

Olive, T. (2012). Working memory in writing. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 485–503). Taylor & Francis. https://doi.org/10.4324/9780203805312.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*, 492–518. http://dx.doi.org/10.1093/applin/24.4.492

Ortega, L. (2009). *Understanding Second Language Acquisition*. London: Hodder Education.

Oushiro, L. (2021). *Introdução à Estatística para Linguistas*. Curso Swirl.

Padilla, A. M., Fan, L., Xu, X., & Silva, D. (2013). Mandarin/English two-way immersion program: Language proficiency and academic achievement. Foreign Language Annals, *46*(4), 661-679. https://doi.org/10.1111/flan.12060

Palmer, M. B. (2006) Using language sampling to analyze adolescent syntactic structures. *Perspectives on School Based Issues, 7*(1), 9-15. https://doi.org/10.1044/sbi7.1.9

Paulston, C. B. (1980). Bilingual Education: Theories and Issues. Newbury House Publishers, Inc., U.S.A. *Canadian Modern Language Review*, *37*(3), 627–628.

Perrotta, G. (2020). The Theory of Mind (ToM): theoretical, neurobiological and clinical profiles. *J Neuroscience and Neurological Surgery*, *7*(1). DOI:10.31579/2578-8868/141

Petitto, L. A., & Kovelman, I. (2003). The bilingual paradox: How signing-speaking bilingual children help us to resolve it and teach us about the brain's mechanisms underlying all language acquisition. *Learning Languages, 8*(3), 5-18. DOI:10.1080/15235882.2013.779618

Petitto, L. A., & Marentette, P. F. (1991). Babbling in the manual mode: Evidence for the ontogeny of language. *Science*, *251*, 1493-6.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied

Linguistics, 15, 36-57.

Poulisse, N. (1997). Language Production in Bilinguals. In A. M. B. de Groot, & J. K. Kroll (Eds.), *Tutorials in bilingualism*: Psycholinguistic perspectives (pp. 201-225). Lawrence Erlbaum.

Pupp Spinassé, K. (2006). Os conceitos de Língua Materna, Segunda Língua e Língua Estrangeira e os falantes de língua alóctones minoritárias no Sul do Brasil. *Revista Congênita, 1*, 01-10.

RSTUDIO TEAM. *RStudio: Integrated Development for R*. RStudio Inc., Boston, MA. http://www.rstudio.com/

Reyes, I. (2012). Biliteracy among children and youths. *Reading Research Quarterly*, *47*(3), 307-327.

Reyes, I. (2006). Exploring connections between emergent biliteracy and bilingualism. *Journal of Early Childhood Literacy*, *6*(3), 267-292.

Rio de Janeiro. Deliberação CEE No 341, de 12 de novembro de 2013. https://www.cpede.com.br/wp-content/uploads/2018/02/delibera%C3%A7%C3%A3o-341-CEE-Estabelece-normas-para-a-oferta-de-Ensino-Bil%C3%ADngue-e-Internacional.pdf.

Rosenbrock, E., Fritzen, M. P., Heinig, O., & Oliveira Martins, L. (2018). (Inter)relações entre práticas de letramentos na escrita de narrativas em alemão e em português por crianças que vivem em contexto de alemão como língua de herança familiar. *Pandaemonium Germanicum, 21*(33), 136-164.

Santa Catarina. Resolução CEE/SC Nº 087, de 22 de novembro de 2016.

Schoonen, R., Van Gelderen, A., Stoel, R. D., Hulstijn, J., & De Glopper, K. (2011). Modeling the development of L1 and EFL writing proficiency of secondary school students. *Language Learning, 61*(1), 31–79. https://doi.org/10.1111/j.1467-9922.2010.00590.x

Schwartz, M., Ibrahim, R., Kahn-Horwitz, J. (2016) Multi-literate experience as a treasure chest for young learners of English as a foreign language. *Reading and Writing: An Interdisciplinary Journal*, *29*(7), 1293-1315.

Soltero-González, L., Sparrow, W., Butvilofsky, S., Escamilla, K., & Hopewell, S. (2016). Effects of a Paired Literacy Program on emerging bilingual children 's biliteracy outcomes in third grade. *Journal of Literacy Research, 48*(1), 80-104.

Soltero-González, L., & Butvilofsky, S. (2016). The early Spanish and English writing development of simultaneous bilingual preschoolers. *Journal of Early Childhood Literacy, 16*(4), 473-497. https://doi.org/10.1177/1468798415602050

Stoller, F. L. (2004). Content-based instruction: perspectives on curriculum planning. *Annual Review of Applied Linguistics*, *24*, 261-283. Cambridge University Press. https://doi.org/10.1017/S0267190504000108

Ströbel, M., Kerz, E., & Wiechmann, D. (2020). The relationship between first and second language writing: Investigating the effects of L1 complexity on L2 complexity in advanced stages

of learning. *Language Learning, 70*(3), 732-767. 10.1111/lang.12394

Sparrow, W., Butvilofsky, S., Escamilla, K., Hopewell, S., & Tolento, T. (2014). Examining the longitudinal biliterate trajectory of emerging bilingual learners in a paired literacy instruction model. *Bilingual Research Journal*, *37*(1), 24-42.

Surrain, S., & Luk, G. (2017). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, 1-15. doi:10.1017/S1366728917000682.

Vaid, J. (2002). Bilingualism. In V. S. Ramachandran (Ed.), *Encyclopedia of the Human Brain*. Academic Press.

Valdés, G. Heritage language students: Profiles and possibilities. (2001) In J. K. Peyton, D. A. Ranard, & S. McGinnis. (Eds.) *Heritage languages in America: Preserving a national resource* (p. 37-77). Center for Applied Linguistics & Delta Systems.

Weigle, S. C. (2005). Second language writing expertise. In K. Johnson (Ed.), *Expertise in second language learning and teaching* (pp. 128-149). Palgrave Macmillan. https://doi.org/10.1057/9780230523470_7.

Weissheimer, J., & Costa, A. C. O. (2022). Biliteracia simultânea na educação bilíngue: efeitos na conectividade da fala e na fluência da leitura de crianças do 5o ano. In: Literacia e Alfabetização: da teoria à prática. (in print).

Williams, C., & Lowrance-Faulhaber, E. (2018). Writing in young bilingual children: Review of research. *Journal of Second Language Writing, 42,* 58-69.

Winter, B. (2019). *Statistics for Linguists: An Introduction Using R.* Routledge. https://doi.org/10.4324/9781315165547

Witkowska, D., Lucas, L., Jelen, M. B., Kin, H., & Norbury, C. (2022). *Development of complex syntax in the narratives of children with English as an Additional Language and their monolingual peers.* (Preprint). https://doi.org/10.31234/osf.io/96dx5

Wong, F., Fillmore, & Valadez, L. (1986). Teaching bilingual learners. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed) (pp. 648-685). Macmillan.

Wright, W. E., & Baker, C. (2017). Key Concepts in Bilingual Education. In: García, O.; Lin, A. M.Y.; May, S. *Bilingual and multilingual education.* (3 ed., pp. 65-80) Springer International Publishing.

Yanden, D., Tsai, T. (2012). Learning how to write in English and Chinese: Young bilingual Kindergarten and first grade children explore the similarities and differences between writing systems. In E. Bauer, & M. Gort. (Eds.) *Early Biliteracy Development: Exploring Young Learners' Use of their Linguistic Resources* (pp. 55-83). New York: Routledge.

Zimmer, M., Finger, I., & Scherer, L. (2008). Do bilinguismo ao multilinguismo: intersecções entre a psicolinguística e a neurolinguística. *Revista Virtual de Estudos da Linguagem:* ReVEL, *6*(11), 1-28.

Zirnstein, M., Bice, K., & Kroll, J. F. (2019). Variation in language experience shapes the consequences of bilingualism. In I. A. Sekerina, L. Spradlin, & V. V. Valian (Eds.*), Bilingualism*

and Executive Function: An Interdisciplinary Approach (pp. 35-48). John Benjamins Publishers.

Appendix 1

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ESTUDOS DA LINGUAGEM
ESPECIALIDADE: PSICOLINGUÍSTICA
Doutoranda: Cristiane Ely Lemke (crisely02@gmail.com)
Doutoranda: Luciana de Souza Brentano (luciana.b@ienh.com.br)
Orientadora: Profa. Dra. Ingrid Finger (ingrid.finger@ufrgs.br)

A/C Coordenação da Unidade Oswaldo Cruz

SOLICITAÇÃO DE AUTORIZAÇÃO DE PESQUISA

NA UNIDADE OSWALDO CRUZ/IENH

O projeto de pesquisa **"Os efeitos da biliteracia em crianças bilíngues brasileiras"** insere-se nos estudos de psicolinguística e tem como objetivo investigar investigar os efeitos do bilinguismo e da alfabetização em duas línguas nos níveis de organização do pensamento e de complexidade sintática a partir de produções orais e escritas em duas línguas (português e inglês) por alunos do Ensino Fundamental inseridos em um contexto de escolarização bilíngue. Através de um estudo longitudinal e transversal, queremos acompanhar o desenvolvimento desses processos cognitivos nas crianças do Ensino Fundamental, fazendo coletas de dados ao longo do ano letivo.

Os participantes deste estudo serão os alunos do 4º ao 7º ano do ensino fundamental. Os pais ou responsáveis pelas crianças deverão autorizar previamente a participação dos seus filhos mediante assinatura de um "Termo de Consentimento Livre e Esclarecido". Os alunos também assinarão um "Termo de Assentimento" caso concordem em participar da pesquisa.

Esta pesquisa é de cunho experimental e serão realizadas tarefas de produção escrita e oral que visam analisar a organização do pensamento e a evolução da escrita tanto em língua inglesa quanto em língua portuguesa. A participação dos alunos e da escola na pesquisa é voluntária e não ocasionará nenhum dano físico ou moral aos participantes ou à escola, sendo a duração das tarefas a única inconveniência.

Dessa forma, solicitamos a autorização da Direção da escola para realizar a coleta de dados da pesquisa descrita acima, mediante a assinatura deste documento.

AUTORIZAÇÃO

Eu, _____, pelo presente documento, declaro que autorizo a realização desta pesquisa no _____ (nome do colégio) pois fui informado, de forma clara e detalhada, livre de qualquer forma de constrangimento e coerção, dos objetivos e justificativa desta pesquisa, dos procedimentos a que os participantes serão submetidos, dos riscos, desconfortos e benefícios e sobre as tarefas que realizaremos, todos acima listados. Fui, igualmente, informado:

· da garantia de receber resposta a qualquer pergunta ou esclarecimento a qualquer dúvida acerca dos procedimentos, riscos, benefícios e outros assuntos relacionados com a pesquisa;

· da liberdade de retirar o consentimento, a qualquer momento, e sem que isso me traga prejuízo;

· da garantia de que os nomes dos alunos, endereço, telefone e nome do colégio serão mantidos em sigilo quando da divulgação dos resultados, e que as informações obtidas serão usadas apenas para fins científicos vinculados ao presente projeto de pesquisa;

· da possibilidade de, após a conclusão da pesquisa, receber uma devolutiva sobre os resultados da pesquisa, se assim desejar.

A pesquisadora responsável por este projeto de pesquisa é a professora Dr. Ingrid Finger (ingrid.finger@ufrgs.br) telefone institucional: 51-3308.6704; endereço institucional: gabinete Nº220 do Prédio Administrativo do Instituto de Letras do Campus do Vale da UFRGS. Quaisquer dúvidas podem ser sanadas junto à doutoranda Cristiane Ely Lemke (crisely02@gmail.com, telefone: 51-991934048) ou junto ao Comitê de Ética em Pesquisa da UFRGS (CEP/UFRGS: 51-33083738). O presente documento será assinado em duas vias de igual teor, uma ficando com o voluntário da pesquisa e outra com a pesquisadora.

Porto Alegre, ___/___/___

_____   _____
Assinatura do Diretor da Escola   Assinatura da responsável pela
coleta de dados

Appendix 2

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ESTUDOS DA LINGUAGEM
ESPECIALIDADE: PSICOLINGUÍSTICA
Doutoranda: Cristiane Ely Lemke (crisely02@gmail.com)
Doutoranda: Luciana de Souza Brentano (luciana.b@ienh.com.br)
Orientadora: Profa. Dra. Ingrid Finger (ingrid.finger@ufrgs.br)

Termo de Consentimento Livre e Esclarecido

Eu, Cristiane Ely Lemke, professora de inglês da IENH e doutoranda em Letras na UFRGS, estou desenvolvendo um estudo sobre bilileracia em situação de bilinguismo de escolha, sob a orientação da Professora Dra. Ingrid Finger. O projeto nos ajudará a acompanhar o desenvolvimento da produção escrita e oral dos alunos nas duas línguas do currículo e poderá auxiliar para uma melhor compreensão desse processo.

Analisaremos duas produções textuais e orais, realizadas durante alguns momentos de aula, em maio e em setembro de 2021 e 2022. As informações que obtivermos serão rigorosamente confidenciais e os nomes dos/das alunos/as serão substituídos por outro em qualquer apresentação ou publicação baseada em nosso estudo.

As tarefas propostas serão adaptadas a partir dos simulados de Cambridge e, desta forma, servirão também como atividades preparatórias para os exames de Cambridge, que ocorrem nos quartos e sextos anos do Ensino Fundamental. Essas atividades serão conduzidas por mim com o auxílio dos professores das turmas.

A participação de seu(sua) filho(a) será de grande valia para nossa pesquisa. Ao aceitar que seu(sua) filho(a) participe, você mantém o direito de dar sua opinião, de fazer perguntas, de desistir de participar da pesquisa, se assim desejar. Se você tem dúvidas ou perguntas, entre em contato comigo pelo e-mail cristiane.l@ienh.com.br. A coordenação da escola tem ciência do trabalho que será realizado.

A pesquisadora responsável por este projeto de pesquisa é a professora Dr. Ingrid Finger (ingrid.finger@ufrgs.br) telefone institucional: 51-3308.6704; endereço institucional: gabinete

Nº220 do Prédio Administrativo do Instituto de Letras do Campus do Vale da UFRGS. Quaisquer dúvidas podem ser sanadas junto à doutoranda Cristiane Ely Lemke (cristiane.l@ienh.com.br, telefone: 51-991934048) ou junto ao Comitê de Ética em Pesquisa da UFRGS (CEP/UFRGS: 51-33083738). O presente documento será assinado em duas vias de igual teor, uma ficando com o voluntário da pesquisa e outra com a pesquisadora.

_____, ___/___/___

_____   _____
Nome do(a) aluno(a)   Assinatura da responsável

        Atenciosamente,

Cristiane Ely Lemke e Ingrid Finger

Appendix 3

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ESTUDOS DA LINGUAGEM
ESPECIALIDADE: PSICOLINGUÍSTICA
Doutoranda: Cristiane Ely Lemke (crisely02@1@gmail.com)
Orientadora: Profa. Dra. Ingrid Finger (ingrid.finger@ufrgs.br)

UFRGS
UNIVERSIDADE FEDERAL
DO RIO GRANDE DO SUL

INSTITUTO
DE LETRAS
U F R G S

Participante nº _____

TERMO DE ASSENTIMENTO

Quero convidar você a participar da pesquisa **"Efeitos do bilinguismo e da biliteracia na organização do pensamento em crianças bilíngues"**. Esse projeto faz parte da tese de doutorado da professora Cristiane Ely Lemke, que tem como objetivo investigar como o uso de duas línguas influencia a cognição de crianças brasileiras. Você está sendo convidado(a) a participar dessa pesquisa porque tem aulas no currículo bilíngue da escola.

Para participar deste estudo, primeiramente o responsável por você deverá autorizar e assinar um termo de consentimento. A sua participação é voluntária, você pode decidir se quer ou não participar da pesquisa. Se você não quiser participar, não vai ter nenhum problema. Mesmo se você aceitar agora, você pode desistir de participar em qualquer momento.

Se você aceitar, você fará atividades, como as da prova de Cambridge,em inglês e em português, durante as aulas de Língua Inglesa. Só as pesquisadoras terão acesso às informações coletadas sobre você e o seu nome não irá aparecer em nenhum lugar. Essas atividades serão avaliadas pela professora da escola, mas ajudarão você nos seus estudos. Depois que a pesquisa acabar, os resultados serão informados em publicações em revistas e apresentações em congressos.

A participação na pesquisa não trará nenhum dano físico ou moral a você. Todas as tarefas serão organizadas de forma a ajudar você na preparação para os exames de Cambridge. A sua participação nesta pesquisa pode nos ajudar a entender melhor como a exposição a mais de uma língua influencia outros processos cognitivos e o desenvolvimento da produção oral e escrita em duas línguas.

Eu, _____, aceito participar desta pesquisa. Sei que a qualquer momento poderei fazer perguntas sobre a pesquisa, e o meu responsável poderá modificar a decisão de participar se quiser.

Novo Hamburgo, ___/___/___

_____   _____
Assinatura do participante   Assinatura da responsável pela coleta de dados

Appendix 4

MATRIZES PROGRESSIVAS COLORIDAS

Escala Especial
Séries A, Ab, B
J. C. RAVEN

FOLHA DE RESPOSTAS

Nome:_____ Sexo:_____   Data de Aplicação:_____

Escola:_____   Data de Nascimento:_____

Série:_____ Período:_____   Idade:_____ a_____ m_____

Examinador:_____ Revisor:_____   Forma de Aplicação:_____

	A			Ab			B	
1			1			1		
2			2			2		
3			3			3		
4			4			4		
5			5			5		
6			6			6		
7			7			7		
8			8			8		
9			9			9		
10			10			10		
11			11			11		
12			12			12		
Σ A			Σ Ab			Σ B		
Consist.			Consist.			Consist.		
Discrep.			Discrep.			Discrep.		

Σ	Percentil	Classif.
Início	Fim	Duração

Obs.:_____

A2

Ав5

B5
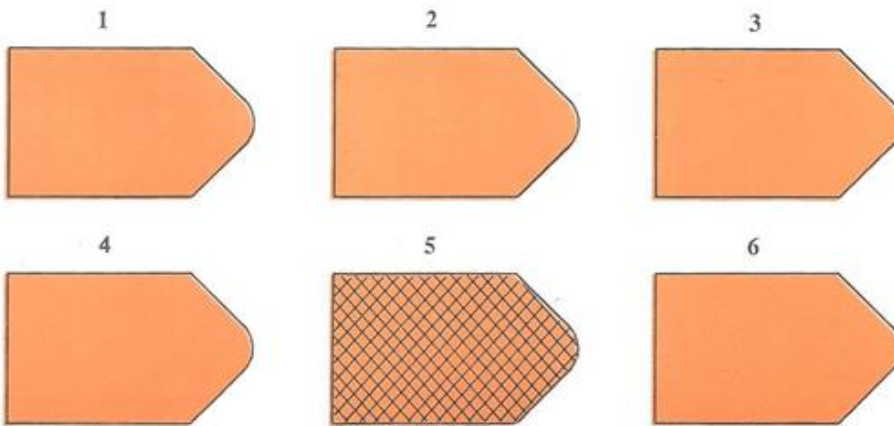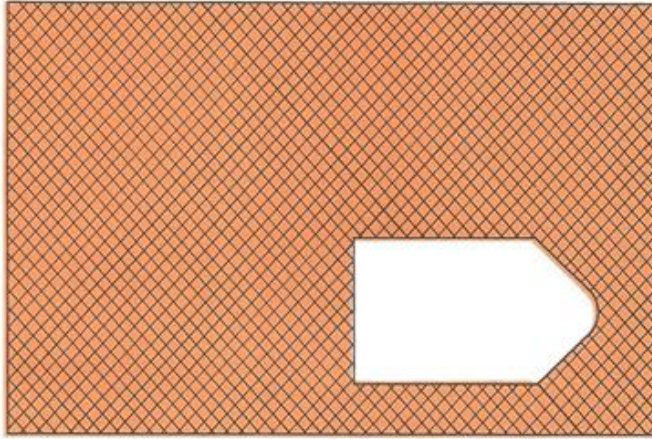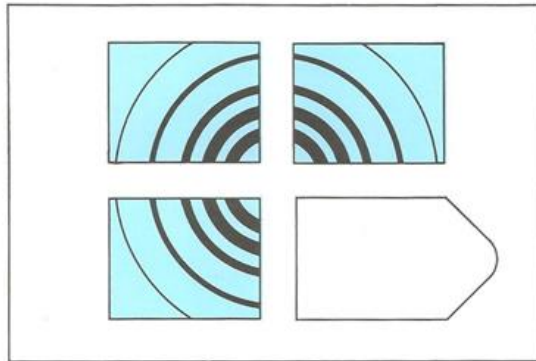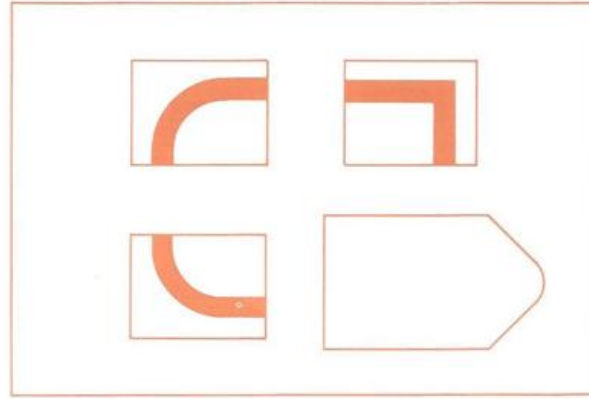
Appendix 5

WRITING ACTIVITY

Look at the five pictures. Write about the story. Write 200 words or more.
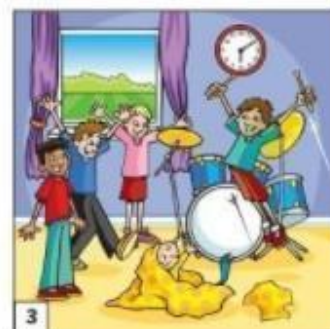
FLYERS SPEAKING. Picture Story

ATIVIDADE DE ESCRITA

Observe as cinco figuras. Escreva sobre esta história. Escreva 200 palavras ou mais.

FLYERS SPEAKING. Picture Story

WRITING ACTIVITY

Look at the five pictures. Write about the story. Write 200 words or more.

Grandma's busy day

Paul   Grandma

ATIVIDADE DE ESCRITA

Observe as cinco figuras. Escreva sobre esta história. Escreva 200 palavras ou mais.

Grandma's busy day

Paul   Grandma

Appendix 6

PROTOCOLO DE TESTE PRESENCIAL

A professora inicia a aula saudando os alunos e explicando que hoje ela irá aplicar uma atividade de escrita com a turma.

Explicar aos alunos que eles não serão avaliados, mas que gostaria que fizessem a escrita do texto de acordo com as instruções e fazendo o seu melhor. Esse trabalho é importante para a análise do desenvolvimento da escrita dos alunos de forma geral, não individual.

Explicar aos alunos que o texto escrito deve ser na língua que receberem a sua atividade. Explicar aos alunos que a história deve ser escrita de acordo com as imagens e que eles podem criar elementos para enriquecer suas histórias, mas que elas precisam seguir a sequência e os fatos das imagens apresentadas. É uma boa ideia primeiro olhar bem as imagens e pensar sobre o que está acontecendo e em como contar essa história.

Explicar que o texto não pode conter diálogos e que terão o tempo de 35 minutos para escrever uma história bem elaborada, como se estivessem contando aquela história para um amigo.

A professora irá compartilhar uma cópia da atividade para cada aluno. Os alunos usarão o chromebook, que estará bloqueado para pesquisa. Explicará que é uma atividade individual e que deve ser realizada em silêncio.

Dizer que irá compartilhar com eles os testes no Google Classroom e que cada um já receberá sua atividade individualmente, na sala de aula da turma. Devem acessar o tópico EXTRA.

Explicar que irá avisá-los quando atingirem o tempo de 30 minutos e que então terão cinco minutos finais para encerrar a atividade.

A professora observa os alunos durante a realização do teste.

Após os 35 minutos, o compartilhamento do documento com os alunos será desfeito.

A professora irá agradecer a participação e envolvimento dos alunos.

Appendix 7

TAREFA ORAL 1

-       Primeiro, observe todas as imagens, elas contam uma história sobre "*Um piquenique com os ursos", "A picnic with the bears".*

-       Você irá me contar essa história EM PORTUGUÊS/EM INGLÊS.

-       No primeiro quadrinho, a gente vê *uma família reunida organizando o carro para sair".*

-       Agora é com você, conte esta história, narrando todos os eventos, usando as imagens como guia.

TAREFA ORAL 2

- Primeiro, observe todas as imagens, elas contam uma história chamada *"Futebol na sala de estar", "Football in the living room".*
- Você irá me contar essa história EM PORTUGUÊS/EM INGLÊS.
- No primeiro quadrinho, a gente vê *o Paul jogando futebol na sala.*
- Agora é com você, conte esta história, narrando todos os eventos, usando as imagens como guia.
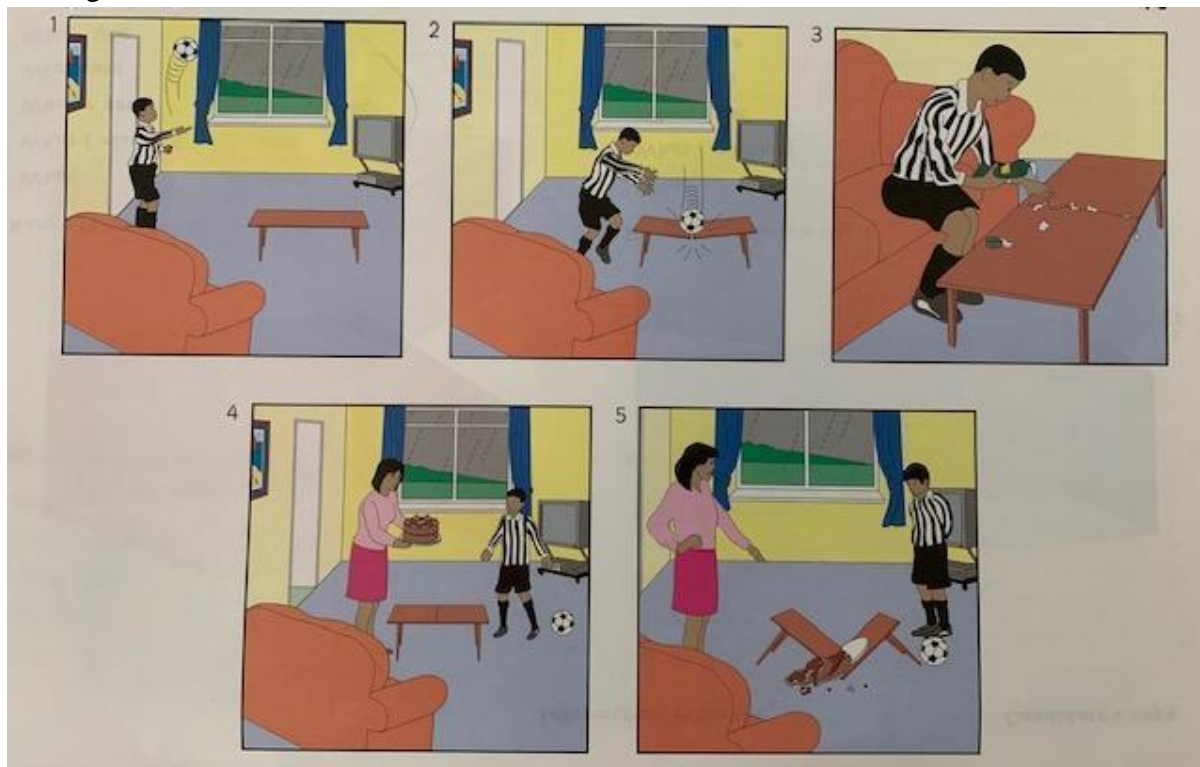
Appendix 8

Table 4: SpeechGraphs attributes and SI regarding Modality

Tests comparing *SpeechGraphs* attributes and SI regarding Modality (Written and Oral)

variables	df	test statistic	p.value	estimate	conf.low	conf.high	method
WRI_WC_L1 and ORAL_WC_L1	-	4545.5	< 0.001	-	-	-	Wilcoxon signed rank test with continuity correction
WRI_WC_L2 and ORAL_WC_L2	-	3693	< 0.001	-	-	-	Wilcoxon signed rank test with continuity correction
WRI_LSC_L1 and ORAL_LSC_L1	-	4532.5	< 0.001	-	-	-	Wilcoxon signed rank test with continuity correction
WRI_LSC_L2 and ORAL_LSC_L2	-	3512.5	< 0.001	-	-	-	Wilcoxon signed rank test with continuity correction
WRI_LCC_L1 and ORAL_LCC_L1	-	4449	< 0.001	-	-	-	Wilcoxon signed rank test with continuity correction
WRI_LCC_L2 and ORAL_LCC_L2	-	3752.5	< 0.001	-	-	-	Wilcoxon signed rank test with continuity correction
WRI_RE_L1 and ORAL_RE_L1	96	5.7	< 0.001	3.68	2.4	4.96	Paired t-test
WRI_RE_L2 and ORAL_RE_L2	-	2683.5	< 0.001	-	-	-	Wilcoxon signed rank test with continuity correction
WRI_PE_L1 and ORAL_PE_L1	96	5.69	< 0.001	3.93	2.56	5.3	Paired t-test
WRI_PE_L2 and ORAL_PE_L2	-	2539	< 0.001	-	-	-	Wilcoxon signed rank test with continuity correction
WRI_IS_L1 and ORAL_IS_L1	101	4.3	< 0.001	0.26	0.14	0.38	Paired t-test
WRI_IS_L2 and ORAL_IS_L2	88	4.74	< 0.001	0.2	0.11	0.28	Paired t-test

Note:
Confidence intervals (CI: conf.low - conf.high) and degrees of freedom (df) were provided for the parametric tests

Source: author

Table 5: SpeechGraphs attributes and SI regarding Language

Tests comparing *SpeechGraphs* attributes and SI regarding Language (Portuguese and English)

variables	df	test statistic	p.value	conf.low	conf.high	method
WRI_WC_L1 and WRI_WC_L2	-	4497.5	< 0.001	-	-	Wilcoxon signed rank test with continuity correction
ORAL_WC_L1 and ORAL_WC_L2	-	3789.5	< 0.001	-	-	Wilcoxon signed rank test with continuity correction
WRI_LSC_L1 and WRI_LSC_L2	99	12.84	< 0.001	23.13	31.59	Paired t-test
ORAL_LSC_L1 and ORAL_LSC_L2	-	4132	< 0.001	-	-	Wilcoxon signed rank test with continuity correction
WRI_LCC_L1 and WRI_LCC_L2	99	12.86	< 0.001	22.61	30.87	Paired t-test
ORAL_LCC_L1 and ORAL_LCC_L2	95	10.58	< 0.001	11.54	16.87	Paired t-test
WRI_RE_L1 and WRI_RE_L2	96	0.14	> 0.05	-1.47	1.69	Paired t-test
ORAL_RE_L1 and ORAL_RE_L2	-	1396.5	< 0.05	-	-	Wilcoxon signed rank test with continuity correction
WRI_PE_L1 and WRI_PE_L2	96	0.31	> 0.05	-1.46	2	Paired t-test
ORAL_PE_L1 and ORAL_PE_L2	-	1310.5	< 0.05	-	-	Wilcoxon signed rank test with continuity correction
WRI_IS_L1 and WRI_IS_L2	99	9.54	< 0.001	0.43	0.66	Paired t-test
ORAL_IS_L1 and ORAL_IS_L2	91	8.67	< 0.001	0.34	0.54	Paired t-test

Note:
Confidence intervals (CI: conf.low - conf.high) and degrees of freedom (df) were provided for the parametric tests

Source: author

Appendix 9

Table 6: Spearman's Correlations between SpeechGraph attributes and SI in oral and written texts

Correlations between *SpeechGraphs* attributes and SI in Written and Oral texts

variables	test statistic	rho	p.value	method
WRI_LSC_L1 and WRI_LSC_L2	54577.09	0.67	< 0.001	Spearman's rank correlation rho
ORAL_LSC_L1 and ORAL_LSC_L2	107435.6	0.27	< 0.01	Spearman's rank correlation rho
WRI_LCC_L1 and WRI_LCC_L2	54747.65	0.67	< 0.001	Spearman's rank correlation rho
ORAL_LCC_L1 and ORAL_LCC_L2	104175.35	0.29	< 0.01	Spearman's rank correlation rho
WRI_RE_L1 and WRI_RE_L2	101242.82	0.33	< 0.001	Spearman's rank correlation rho
ORAL_RE_L1 and ORAL_RE_L2	106515.19	0.25	< 0.05	Spearman's rank correlation rho
WRI_PE_L1 and WRI_PE_L2	106061.54	0.3	< 0.01	Spearman's rank correlation rho
ORAL_PE_L1 and ORAL_PE_L2	103996.79	0.27	< 0.01	Spearman's rank correlation rho
WRI_IS_L1 and WRI_IS_L2	143546.37	0.14	> 0.05	Spearman's rank correlation rho
ORAL_IS_L1 and ORAL_IS_L2	127056.32	0.02	> 0.05	Spearman's rank correlation rho

Source: author

Table 7: Spearman's Correlations between SpeechGraph attributes and SI in Portuguese and English texts

Correlations between *SpeechGraphs* attributes and SI in Portuguese and English texts

variables	rho	test statistic	p.value	method
WRI_LSC_L1 and WRI_IS_L1	0.36	127718.08	< 0.001	Spearman's rank correlation rho
WRI_LSC_L2 and WRI_IS_L2	0.35	117865.82	< 0.001	Spearman's rank correlation rho
ORAL_LSC_L1 and ORAL_IS_L1	0.36	120069.3	< 0.001	Spearman's rank correlation rho
ORAL_LSC_L2 and ORAL_IS_L2	0.28	102242.49	< 0.01	Spearman's rank correlation rho
WRI_LCC_L1 and WRI_IS_L1_b	0.35	128778.82	< 0.001	Spearman's rank correlation rho
WRI_LCC_L2 and WRI_IS_L2_b	0.36	116969.01	< 0.001	Spearman's rank correlation rho
ORAL_LCC_L1 and ORAL_IS_L1_b	0.35	121259.31	< 0.001	Spearman's rank correlation rho
ORAL_LCC_L2 and ORAL_IS_L2_b	0.27	103814.76	< 0.01	Spearman's rank correlation rho
WRI_RE_L1 and WRI_IS_L1_c	0.19	147070.5	> 0.05	Spearman's rank correlation rho
WRI_RE_L2 and WRI_IS_L2_c	0.29	130030.05	< 0.01	Spearman's rank correlation rho
ORAL_RE_L1 and ORAL_IS_L1_c	0	187447.92	> 0.05	Spearman's rank correlation rho
ORAL_RE_L2 and ORAL_IS_L2_c	0.24	105047.13	< 0.05	Spearman's rank correlation rho
WRI_PE_L1 and WRI_IS_L1_d	0.18	149915.83	> 0.05	Spearman's rank correlation rho
WRI_PE_L2 and WRI_IS_L2_d	0.27	132543.31	< 0.01	Spearman's rank correlation rho
ORAL_PE_L1 and ORAL_IS_L1_d	0.07	174794.8	> 0.05	Spearman's rank correlation rho
ORAL_PE_L2 and ORAL_IS_L2_d	0.24	105143.94	< 0.05	Spearman's rank correlation rho

Source: author

Table 8: Spearman's Correlations between SpeechGraph attributes and SI in English texts and English Proficiency

Correlations between *SpeechGraphs* attributes and SI in English texts and English Proficiency

variables	rho	parameter	conf.low	conf.high	test statistic	p.value	method
WRI_LSC_L1 and L2 Proficiency	0.3	-	-	-	55049.55	< 0.01	Spearman's rank correlation rho
WRI_LSC_L2 and L2 Proficiency	0.47	-	-	-	45262.86	< 0.001	Spearman's rank correlation rho
ORAL_LSC_L1 and L2 Proficiency	0.18	74	-0.05	0.39	1.54	> 0.05	Pearson's product-moment correlation
ORAL_LSC_L2 and L2 Proficiency	0.38	-	-	-	32260.93	< 0.01	Spearman's rank correlation rho
WRI_LCC_L1 and L2 Proficiency	0.29	-	-	-	55884.81	< 0.01	Spearman's rank correlation rho
WRI_LCC_L2 and L2 Proficiency	0.5	-	-	-	42242.12	< 0.001	Spearman's rank correlation rho
ORAL_LCC_L1 and L2 Proficiency	0.21	74	-0.02	0.41	1.82	> 0.05	Pearson's product-moment correlation
ORAL_LCC_L2 and L2 Proficiency	0.41	-	-	-	30660.14	< 0.001	Spearman's rank correlation rho
WRI_RE_L1 and L2 Proficiency	0.33	-	-	-	49056.43	< 0.01	Spearman's rank correlation rho
WRI_RE_L2 and L2 Proficiency	0.3	-	-	-	59470.71	< 0.01	Spearman's rank correlation rho
ORAL_RE_L1 and L2 Proficiency	0.2	-	-	-	58694.09	> 0.05	Spearman's rank correlation rho
ORAL_RE_L2 and L2 Proficiency	0.24	-	-	-	39595.1	< 0.05	Spearman's rank correlation rho
WRI_PE_L1 and L2 Proficiency	0.34	-	-	-	48422.89	< 0.01	Spearman's rank correlation rho
WRI_PE_L2 and L2 Proficiency	0.33	-	-	-	57125.76	< 0.01	Spearman's rank correlation rho
ORAL_PE_L1 and L2 Proficiency	0.14	-	-	-	62813.77	> 0.05	Spearman's rank correlation rho
ORAL_PE_L2 and L2 Proficiency	0.23	-	-	-	40112.29	> 0.05	Spearman's rank correlation rho
WRI_IS_L1 and L2 Proficiency	-0.05	-	-	-	86651.18	> 0.05	Spearman's rank correlation rho
WRI_IS_L2 and L2 Proficiency	0.27	-	-	-	57981.99	< 0.05	Spearman's rank correlation rho
ORAL_IS_L1 and L2 Proficiency	-0.2	-	-	-	91218.33	> 0.05	Spearman's rank correlation rho
ORAL_IS_L2 and L2 Proficiency	0.23	-	-	-	41877.98	> 0.05	Spearman's rank correlation rho

Source: author

Table 9: Spearman's Correlations between SpeechGraph attributes and SI between written and oral texts

Correlations between *SpeechGraphs* attributes and SI between Written and Oral texts

variables	test statistic	rho	p.value	method
WRI_WC_L1 and ORAL_WC_L1	122802.58	0.26	< 0.001	Spearman's rank correlation rho
WRI_WC_L2 and ORAL_WC_L2	57198.76	0.54	< 0.001	Spearman's rank correlation rho
WRI_LSC_L1 and ORAL_LSC_L1	116251.02	0.3	< 0.001	Spearman's rank correlation rho
WRI_LSC_L2 and ORAL_LSC_L2	50201.65	0.6	< 0.001	Spearman's rank correlation rho
WRI_LCC_L1 and ORAL_LCC_L1	120791.76	0.28	< 0.001	Spearman's rank correlation rho
WRI_LCC_L2 and ORAL_LCC_L2	42759.53	0.66	< 0.001	Spearman's rank correlation rho
WRI_RE_L1 and ORAL_RE_L1	125740.76	0.17	< 0.001	Spearman's rank correlation rho
WRI_RE_L2 and ORAL_RE_L2	96368.68	0.21	< 0.001	Spearman's rank correlation rho
WRI_PE_L1 and ORAL_PE_L1	123379.56	0.19	< 0.001	Spearman's rank correlation rho
WRI_PE_L2 and ORAL_PE_L2	95122.48	0.22	< 0.001	Spearman's rank correlation rho
WRI_IS_L1 and ORAL_IS_L1	161228.12	0.09	< 0.001	Spearman's rank correlation rho
WRI_IS_L2 and ORAL_IS_L2	91658.4	0.22	< 0.001	Spearman's rank correlation rho

Source: author