UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSITCOS

MARA PASSOS GUIMARÃES

STRUCTURAL PERSISTENCE AND SURPRISAL: IMPLICATIONS FOR PROFICIENCY-MODULATED DISTRIBUTIONAL LEARNING IN LATE BILINGUALS

BELO HORIZONTE
2018

MARA PASSOS GUIMARÃES

STRUCTURAL PERSISTENCE AND SURPRISAL: IMPLICATIONS FOR PROFICIENCY-MODULATED DISTRIBUTIONAL LEARNING IN LATE BILINGUALS

Tese de doutorado apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais como parte do requisito para obtenção do título de doutora em Estudos Linguísticos.

Área de concentração: Linguística Teórica e Descritiva
Linha de pesquisa: Processamento da Linguagem
Orientador: Prof. Dr. Ricardo Augusto de Souza

BELO HORIZONTE
2018

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca FALE/UFMG

Guimarães, Mara Passos.

G963s

Sructural persistence and surprisal [manuscrito] : implications for proficiency-modulated distributional learning in late bilinguals / Mara Passos Guimarães. – 2018.

92 f., enc. : il., tab., grafs., color.

Orientador: Ricardo Augusto de Souza.

Área de concentração: Lingüística Teórica e Descritiva.

Linha de pesquisa: Processamento da Linguagem.

1. Língua inglesa – Estudo e ensino – Falantes estrangeiros – Teses. 2. Aquisição da segunda linguagem – Teses. 3. Bílinguismo – Teses. 4. Psicolínguistica. Souza, Ricardo Augusto de. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD : 420.7

CRB-6/2616

UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

FOLHA DE APROVAÇÃO

Structural persistence and surprisal: implications for proficiency-modulated distributional learning in late bilinguals

MARA PASSOS GUIMARÃES

Tese submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGUÍSTICOS, como requisito para obtenção do grau de Doutor em ESTUDOS LINGUÍSTICOS, área de concentração LINGUÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Processamento da Linguagem.

Aprovada em 10 de agosto de 2018, pela banca constituída pelos membros:

Prof(a). Ricardo Augusto de Souza - Orientador
UFMG

Prof(a). Cândido Samuel Fonseca de Oliveira
CEFET-MG

Prof(a). Marisa Mendonça Carneiro
UFMG

Prof(a). Larissa Santos Ciríaco
UFMG

Prof(a). Mahayana Cristina Godoy
UFRN

Belo Horizonte, 10 de agosto de 2018.

To Giovana and Iara (I know, this isn't nearly as good as cake)

ACKNOWLEDGEMENTS

To Dr. Ricardo de Souza, for the unwavering support and guidance;

A minha família, por priorizar a minha urgência ao seu planejamento;

A Marinela, por me emprestar a voz, o teto, e o carinho, tão essenciais para a conclusão deste trabalho;

A Thaís e Mahayana, pela presença em vários momentos de dúvida;

A Diana, pela preciosa companhia e amizade na fase potiguar deste processo;

A Renata, pela dosagem ideal entre afeto e profissionalismo dos quais esta tese é prova;

E aos meus amigos, que diariamente me lembram a pessoa de sorte que eu sou.

“It is no coincidence that “aspiration” means both hope and the act of breathing.”

Ted Chiang

RESUMO

Entende-se que bilíngues de alta proficiência compartilham representações estruturais abstratas entre a primeira e a segunda língua – L1 e L2, respectivamente (Hartsuiker et al., 2004; Bernolet et al., 2013; Guimarães, 2016; Souza et al., 2014, Souza and Oliveira, 2014). Compartilhamento representacional é modulado por proficiência na L2 (Bernolet et al., 2013), construto que se baseia em níveis de automaticidade linguística de *loci* de memória (conhecimento implícito ou explícito) para definir e mensurar linguagem (Ullman, 2004; Hustijn, 2015). O propósito deste estudo é investigar se alta proficiência em L2 implica que mecanismos subjacentes de previsão de erro e abstração de regras gramaticais baseadas em aprendizado distribucional (*distributional learning*) são estendidos ao processamento da L1 por bilíngues tardios. Aprendizado distribucional diz respeito ao processo de derivação de generalizações abstratas acerca da linguagem a partir de pistas estatísticas – particularmente a distribuição de frequências de um determinado aspecto da língua. Os estudos desta tese foram elaborados para responder questões sobre bilíngues cuja L1 é o BP e a L2 é o inglês sob duas teorias concorrentes de aprendizado de linguagem: a teoria de ativação lexical (baseada no modelo de ativação residual de Malhotra et al., 2008) e a teoria de aprendizado implícito (baseada no modelo bifurcado proposto por Chang et al., 2006). Estudos 1 e 2 foram análises de corpus sobre sensibilidade a efeitos de *surprisal* e cumulatividade de passivas no corpus do BP falado C-Oral-Brasil I, cujo objetivo foi oferecer estimativas de persistência sintática em BP correspondentes àquelas do inglês em Jaeger e Snider (2007). O estudo 3 buscou investigar diferenças na sensibilidade a efeitos de *surprisal* e cumulatividade em bilíngues e monolíngues. Os resultados dos estudos 1 e 2 indicam que os dados do BP não apresentaram efeitos de *priming* estrutural. O estudo 3 mostrou que há uma tendência crescente dos efeitos de *priming* em monolíngues, levando à conjectura de que existe um mecanismo de aprendizado distribucional subjacente à L1 e à L2, modulado por proficiência e similaridade estrutural.

ABSTRACT

High-proficiency bilinguals are believed to share abstract structural representations between the first and the second language – L1 and L2, respectively (Hartsuiker et al., 2004; Bernolet et al., 2013; Guimarães, 2016; Souza et al., 2014, Souza and Oliveira, 2014). Representational sharing is believed to be modulated by L2 proficiency (Bernolet et al., 2013), construct that relies on levels of language automaticity and loci of memory (implicit or explicit knowledge) to define and measure knowledge of language (Ullman, 2004; Hustijn, 2015). The purpose of this study is to investigate whether high L2 proficiency entails that underlying mechanisms of prediction error and rule abstraction based on distributional learning extends over to L1 processing by late bilinguals. Distributional learning refers to the process of deriving abstract generalizations about language from statistical cues – particularly, the frequency distributions of a given aspect of language. The studies in this dissertation were designed to answer questions about L1 Brazilian Portuguese L2 English bilinguals under two competing theories of language learning: a lexical activation account (based on the trailing-activation model proposed by Malhotra et al., 2008) and an implicit learning account (based on the dual-path model proposed by Chang et al., 2006), which differ in their predictions about properties of syntactic persistence (Bock, 1986). Studies 1 and 2 were corpus analyses of passive surprisal-sensitivity and cumulativity of passives in the corpus of spoken BP C-Oral-Brasil I, that aimed to provide syntactic persistence estimates in BP analogous to those of English provided by Jaeger and Snider (2007). Study 3 observed differences in surprisal-sensitivity and cumulativity effects on bilinguals and monolinguals. Results from studies 1 and 2 indicate that the data set for BP did not present priming effects. Study 3 showed an ascending tendency in priming effects on monolinguals, which led to the conjuncture that there is a mechanism of distributional learning that underlies L1 and L2 alike, modulated by proficiency and structural similarity.

LIST OF FIGURES

Figure 1 - Levelt et al. (1999) lexical access network 22
Figure 2 - Model of bilingual sentence production by Hartsuiker et al. (2004) 23
Figure 3 - Trailing-activation network 32
Figure 4 - Dual-path model 34
Figure 5 - Picture used to describe the event "arrest" 35
Figure 6 - Prime surprisal based on prime verb's passive bias (Jaeger and Snider, 2007) 42
Figure 7 - Cumulativity in passives (Jaeger and Snider, 2007) 44
Figure 8 - Individual correlations on structure choice (passive surprisal) 47
Figure 9 - Individual correlations on structure choice (passive cumulativity) 53
Figure 10 - Interaction between passives produced and target verb bias (passive cumulativity) 55
Figure 11 - Image used in the event "push" 61
Figure 12 - Image used in the event "mug" 61
Figure 13 - Image used in the event "run" 61
Figure 14 - Image used in the event "show" 62
Figure 15 - Image used in the event "fan" 64
Figure 16 - Effects of lexical identity on structure choice 65
Figure 17 - Interaction between prime type and lexical identity 66
Figure 18 - Production of passives by linguistic profile 67
Figure 19 – Interaction between prime type and linguistic profile 68
Figure 20 - Passive cumulativity on choice of structure 68
Figure 21 - Image used in the event "chase" 71
Figure 22 - Image portraying the event "strike" 74

LIST OF TABLES

Table 1 – Summary of passive surprisal analysis (Jaeger and Snider, 2007) 42
Table 2 - Summary of passive cumulativity analysis (Jaeger and Snider, 2007).....43
Table 3 - Summary of passive surprisal in C-Oral-Brasil I.....47
Table 4 - Frequencies of passives in C-Oral-Brasil I and Penn Treebank corpora....50
Table 5 - Statistical description of verb biases in BP and English 51
Table 6 - Summary of passive cumulativity in C-Oral-Brasil I.....53
Table 7 - Structures used in voice alternation descriptions 63
Table 8 - Interaction between choice of verb and profile.....66
Table 9 - Interaction between prime type and profile on choice of structure 67
Table 10 - Production in free and primed tasks.....69

LIST OF APPENDICES

1. APPENDIX 1: verb passive biases from C-Oral-Brasil I (Raso and Mello, 2012)...88

2. APPENDIX 2: verb passive biases from SBCSAE (Du Bois et al., 2000-2005)….91

TABLE OF CONTENTS

1. THEORETICAL BACKGROUND 19

1.1. Summary of Guimarães (2016) 19
1.2. Delimiting the construct of L2 proficiency 20
1.3. Bilingual shared representations as a function of L2 proficiency 25
1.4. Distributional learning 27
1.5. Structural priming 29
1.6. Activation-based and implicit learning accounts of language learning 30
1.7. Surprisal and cumulativity 35
1.8. Syntax of oral production 36

2. METHODOLOGY 40

2.1. Jaeger and Snider (2007) 41
2.2. Study 1: surprisal-sensitivity of passives in BP 44
2.2.1. Data 45
2.2.2. Method 45
2.2.3. Analysis 46
2.2.4. Results 48
2.3. Study 2: cumulativity 52
2.3.1. Data 52
2.3.2. Method 52
2.3.3. Analysis 52
2.3.4. Results 54
2.4. Surprisal-sensitivity and cumulativity in the C-Oral-Brasil I corpus: discussion 55
2.5. Study 3 57
2.5.1. Design: contributions from Bock (1986) and Guimarães (2016) 57
2.5.2. Predictions 58
2.5.3. Participants 59
2.5.4. Material 60
2.5.5. Procedures 62
2.5.6. Voice alternation data 62
2.5.7. Results 64
2.5.8. Discussion 68

3. GENERAL DISCUSSION 73

3.1. Structural priming as learning....73
3.2. Distributional learning in late bilingualism....74
3.3. Similarity modulation on shared representations between L1 and L2 77
3.4. Late L2 learning and processing as byproducts of surprisal....81
4. REFERENCES 83
5. APPENDIX 1: verb passive biases from C-Oral-Brasil I (Raso and Mello, 2012)88
6. APPENDIX 2: verb passive biases from SBCSAE (Du Bois et al., 2000-2005) ..91

INITIAL CONSIDERATIONS

It is now widely accepted that bilinguals are not two monolinguals in one mind, but speakers with a distinct linguistic system that shares representations at some level. Thus, psycholinguistics research on bilingualism is mainly concerned about how these representations are related in memory. The nature or extent of such sharing has been under investigation by a number of psycholinguistic studies (Hartsuiker et al., 2004; Souza et al., 2014; among many others), and comprehensive models of bilingual sentence comprehension and production have been proposed (Djisktra and Van Heuven, 2002; Hartsuiker et al., 2004; among others). A point of convergence is that it is not yet possible to generalize findings of shared representation over constructions, languages, or bilingual profiles.

The present study is a development of the findings reported by Guimarães and Souza (2016) and Guimarães (2016) concerning the passive structure in Brazilian Portuguese (henceforth BP) and L1 BP speakers' behavior towards it. The discrepancy between the high levels of acceptance of the passives and the extremely few instances of its production by BP monolinguals reported by Guimarães (2016) gives support to the idea that production and comprehension are two different but related processes that show different levels of sensitivity to frequency effects. Although findings in studies about L2 English interference on L1 BP comprehension have provided substantial evidence of shared representations in late bilinguals (Souza et al., 2014; Souza and Oliveira, 2014), there is still a need for research addressing the issue of such influence on L1 oral production. The present study thus relies on spoken language, both compiled in corpora (Raso and Mello, 2012) and elicited in laboratory, to investigate further whether and to what extent exposure to L2 English changes production patterns in L1 BP speakers.

In addition to the contribution to literature about this particular linguistic pair in oral production, this dissertation aims to open the discussion concerning L2 proficiency as an indicator of the L2 sharing underlying mechanisms of distributional learning and error prediction with the L1, and possibly with general cognitive processes. Bilinguals have been observed to behave similarly to native speakers in relation to both acceptability of unlicensed structures from the L2 in the L1 and increased production of infrequent structures in L1 due to distributional properties of the L2 (Souza et al.,

2014; Souza and Oliveira, 2014; Guimarães, 2016). It is important to note that the similarity here mentioned is not related to notions of native-likeness or ultimate attainment (as defined in Ortega, 2009), since L1 and L2 are not considered separate systems. What is being argued is that the bilingual's familiarity and automaticity concerning structures that show different properties between L1 BP and L2 English differ from monolinguals' and approximates to the L2 native speakers'. Consequently, L1 and L2 are supposed to converge learning and processing mechanisms.

The construct underlying all theories and assumptions of the present study is frequency and its effects on perception, production, and prediction. Implicit learning accounts of language learning (as all error-based accounts) posit that language processing necessarily entails language learning, and it follows that the more episodes of linguistic processing a speaker experiences, the more statistical adjustments and abstract generalizations the speaker is able to make (Chang et al., 2006; Jaeger and Snider, 2007). Such generalizations are byproducts of statistical and distributional learning, which is the process of acquiring information about distributions of elements in the language that determine the constraints or grammatical rules of that language.

The passive has been chosen as the target construction based on characteristics that may be informative to the bilingualism effects under investigation in this study. First, the passive is syntactically and morphologically identical in BP and in English, presenting a promoted object, a copula verb, the main verb in the participle form and an optional agentive by-phrase. This provides a baseline that allows us to compare the structure in both languages in aspects other than its surface form, such as information and prosodic structure, event semantics, and even pragmatic constraints. Second, the passive has been used as a target construction in a number of studies (e.g. Bock, 1986; Pickering and Branigan, 1998; Bock and Griffin, 2000; Hartsuiker et al., 2004; Jaeger and Snider, 2007; Jaeger and Snider, 2013), offering data from other languages to which it will be possible to compare our results. Finally, the discrepancy in the production of passives between BP and English speakers shown by Guimarães and Souza (2016) provided the starting point for the analysis of surprisal and structural priming effects, which are the perspective under which we will analyze bilingualism phenomena in this study.

It is important to highlight that the passive is not taken as a byproduct of transformational processes, but as a construction according to Goldberg (1995): an

independent theoretical entity represented in the procedural memory of the speaker (Goldberg, 1995; Ellis, 2003). Thus, the meaning of propositions in the passive does not depend solely on the lexical items occurring in them, but are instead a combination of the prototypical meaning of the construction and the semantic properties of the verb. Particularly, the passive is considered a complex construction that relates directly to the speaker's pragmatic knowledge and is motivated by the perception and categorization of the world. In Ellis's words,

> "(…) what we express reflects which parts of an event attract our attention; depending on how we direct our attention, we can select and highlight different aspects of the frame, thus arriving at different linguistic expressions" (Ellis, 2003; p. 65).

Jaeger and Snider (2007) define structural priming as "the phenomenon that a structure's a posteriori probability of occurring is increased – after another instance of the same structure – compared to its a priori probability" (p. 26). We analyze structural priming effects in the light of error-based accounts of learning, which argue that both structural priming and language learning share underlying mechanisms (namely implicit learning). Occurrences of cross-linguistic structural priming – that is, a structure in one language priming its subsequent use in the other – are taken as evidence of bilingual shared representations (e.g. Hartsuiker et al., 2004; Dussias and Sagarra, 2007; Bernolet et al., 2013). Therefore, the main hypothesis of this study is that bilinguals share representations to a level that linguistic episodes in either L1 or L2 cause the linguistic system to adapt its expectations (i.e. learn from prediction error) in both languages. There has been little data concerning late bilinguals' abilities to predict L1 and L2 based on one general mechanism of distributional learning, with the majority of studies focusing on first language acquisition by infants. This hypothesis extrapolates the theory about distributional learning from children to adults, and from first to second language learning, based on the shared construct of frequency underlying the (related) mechanisms of learning and prediction. We expect that it will be possible to understand whether there needs to be a point of language ability from which bilinguals will entirely share expectation adaptations and language distributional properties. This study aims to contribute to a better understanding of language processing mechanisms of late bilinguals whose L1 is BP and whose L2 is English, in a context of L1 dominance, specifically in terms of L2 influences on L1 oral production of the passive structure.

Chapter 2 presents the theoretical approaches underlying this study, and chapter 3 outlines the methodology employed to investigate distributional learning in bilinguals. Finally, chapter 4 provides an overview of the findings and a discussion of the implications of structural priming (and its properties of surprisal and cumulativity) on the hypothesis that late bilinguals employ mechanisms of distributional learning in the L1 and the L2 alike.

1. THEORETICAL BACKGROUND

1.1. *Summary of Guimarães (2016)*

In an analysis of L2 effects on the L1, Guimarães (2016) conducted a sentence elicitation study in which BP monolinguals and high-proficiency L1 BP L2 English bilinguals were instructed to describe images that depicted transitive events. The author observed that bilinguals produced significantly more passives than did monolinguals in the oral task (p = 0.017), following the tendency of passives to occur more frequently in English than in BP (Guimarães and Souza, 2016).

The results in Guimarães (2016) have been attributed to a reconfiguration of grammatical restrictions in the bilingual's mind, to better accommodate constructions learned from the L2 (cf. Souza, Oliveira, Guimarães and Almeida, 2014). Souza et al. (2014) propose such a reconfiguration from observation of bilinguals' higher acceptance of unlicensed constructions learned from L2 English in the L1, such as the caused motion alternation (sentence 1) and the adjectival resultative (sentence 2), in relation to monolinguals' acceptance levels[1]:

1. * ***O instrutor correu os meninos pelo parque.***
   The instructor ran the boys around the park.
   'The instructor ran the boys around the park.'
2. * ***O garçom arrumou a mesa e a esfregou limpa.***
   The waiter set the table and it-obl wiped clean.
   'The waiter set the table and wiped it clean.'

Since the passive structure is licensed in both BP and English, the results from Guimarães (2016) suggest that the reconfiguration of grammatical restrictions observed in Souza et al. (2014) were expanded to semantic-pragmatic constraints on the passive in BP. Both studies by Guimarães (2016) and Souza et al. (2014) rely largely on the assumption that bilinguals' linguistic systems share representations between the L1 and the L2 (Grosjean, 1989; Hartsuiker et al., 2004).

The frequency effects found in Guimarães (2016) give support to the model of bilingual sentence production by Hartsuiker et al. (2004). Since the task did not include any type of structural manipulation and was conducted solely in L1 BP, the difference in passive productivity between the two experimental groups is a result of recalibrated

[1] However, see section 3.3 for a discussion about results in Spanish by Trujillo (2018).

distributions of the construction from L2 experience. Note that, in this scenario, L2 experience translates as L2 proficiency.

1.2. *Delimiting the construct of L2 proficiency*

L2 proficiency is a troublesome but fundamental construct for psycholinguistics of bilingualism. A comprehensive definition of L2 proficiency depends on defining the nature of the knowledge involved, and there has been some debate over what linguistic cognition entails. Early models (Lado, 1961; Carroll, 1972; in Hustijn, 2015) defined language proficiency on a two-dimensional axis of components of linguistic knowledge (syntax, morphology, etc.) and the four language skills (reading, listening, speaking, and writing), but failed to include a situational context – i.e. the more "peripheral" skills of language use and communication (Hulstjin, 2015). Most models of proficiency were based on a distinction analogous to Chomsky's notions of linguistic competence and performance (1965): Hymes (1972) divided proficiency into knowledge of the language system and knowledge of the communicative situation; Canale and Swain (1980; in Hustijn, 2015) proposed a model divided into grammatical, sociolinguistic, and strategic knowledge. Interestingly, Canale and Swain's (1980) account of proficiency also presented a subcomponent of probability rules that, according to Hustijn (2015), "has received little attention in the literature; it appears remarkably modern when viewed from a current usage-based network" (p. 39). The relation between L2 proficiency and knowledge of frequency will be resumed in chapter 4.

From a general cognitive perspective, Ullman (2004) points out that the fact that the brain performs "computations on different domains of information" (i.e., it is topographically organized) suggests that "analogous computations may underlie a range of cognitive domains, including language" (p. 232). According to his declarative/procedural model, the systems that underlie declarative memory (knowledge about facts and events) are the same that underlie the mental lexicon; analogously, the systems that underlie procedural memory (implicit memory system) also underlie the mental grammar. Although the interface between declarative and procedural memory is under some debate (Soares-Silva, 2016), the process of memory storage is directly related to the frequency distribution of the linguistic expression (Ullman, 2004; p. 245).

Declarative and procedural memory can be thought of as analogous to controlled and automatic domain-general cognitive processes (Schneider and Shiffrin, 1977). Automatic processes are sets of associated memory nodes that rely very lightly (or not at all) on working memory, meaning that the activation of a node activates the connected nodes without necessarily demanding explicit attention or any control by the subject. Conversely, controlled processes depend heavily on working memory capacity and require explicit attention and control. Once again, the automaticity of a process takes place through the frequency of activation of stimuli and response (Schneider and Shiffrin, 1977), which brings us to the matter of the role of automaticity in language proficiency.

Segalowitz and Hulstijn (2005) illustrate the role of automaticity in L1 with an analysis of its scope on the lexical access model of sentence production proposed by Level et al. (1999, figure 1). The lexical network underlying lexical access is a feedforward system subdivided into a conceptual stratum, a lemma stratum, and a form stratum. Nodes in the conceptual level represent lexical concepts, which are sensitive to perceptual (namely visual and auditory) input. Once the lexical concept is chosen, it activates semantically related lemma nodes and their combinatorial possibilities in the lemma stratum in order to select the appropriate lemma. In turn, the lemma is morphophonologically encoded in the form stratum, producing a morpheme that is then phonetically encoded. The output of the phonetic encoding is the appropriate articulatory gestures for the word to be executed by the articulatory system. In parallel, the speech comprehension system monitors the output of the speech production system to identify errors, disfluencies and other delivery issues.

Departing from Kahneman's (1973) definition of automaticity as "the absence of attentional control in the execution of a cognitive activity" (p. 371), Segalowitz and Hustijn (2015) affirm that the processes of grammatical and phonological encoding (in the lemma and the form strata, respectively), as well as articulation and self-monitoring, are largely automatic due to the modularity that underlies these processes. Lexical selection is a statistical mechanism that favors the highest activated lemma and, consequently, nodes for grammatical and morphophonological encoding as well as gestural scores for articulation are activated sequentially. The process of lexical concept choice taking place in the conceptual stratum, on the other hand, is not considered to be automatic. It is an attention-based process that occurs during the

unfolding of the communicative event and concept preparation[2], and does not rely on modular activation (as do the remaining processes of the lexical network).

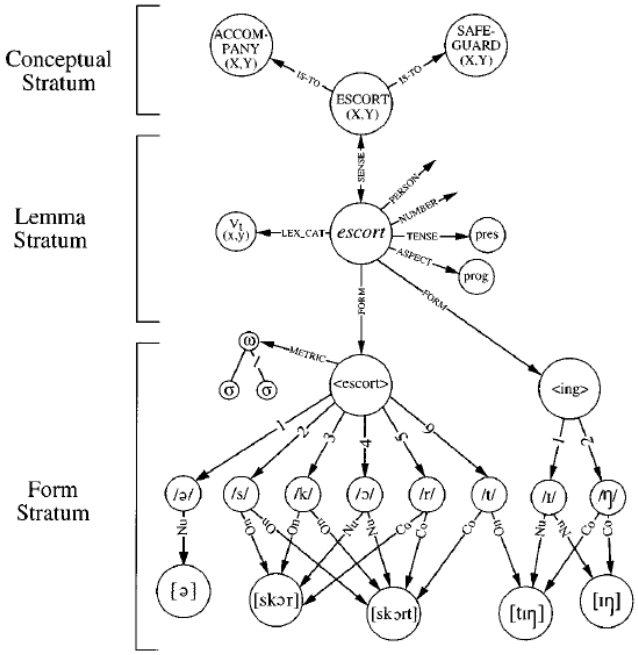

Figure 1 - Levelt et al. (1999) lexical access network

Based on the model of lexical access by Levelt et al. (1999), Hartsuiker et al. (2004) proposed a model of bilingual sentence production that presents the same modular lexical network (figure 2): the conceptual, lemma, and form strata interact in a feedforward system where the output from the conceptual stratum motivates activation of nodes in the lemma stratum, which, in turn, provides input for the selection of the target language word-form. Lemma nodes in this model are unspecified for language: language selection takes place in the form stratum, as language “tags” are connected to the lexical items selected for the activated structure.

[2] Process leading up to the activation of a lexical concept (Level et al., 1999; p. 3).

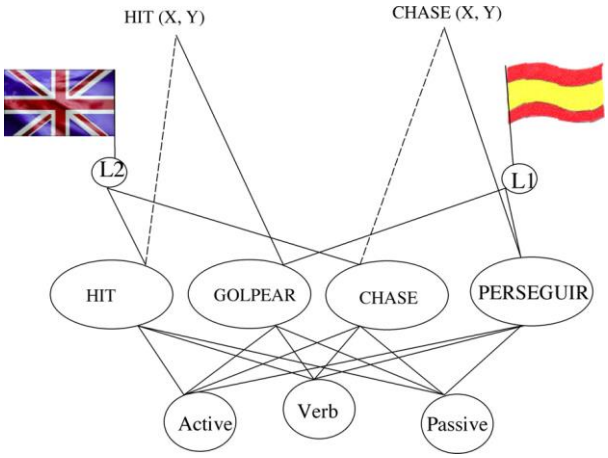

Figure 2 - Model of bilingual sentence production by Hartsuiker et al. (2004)

The locus of language selection in bilingual speech production has been a point of contention within psycholinguistics of bilingualism, with direct and important consequences to the scope of automaticity in L2 proficiency and, consequently, representational sharing. Language-selective models of bilingual speech production assume that the intention of speaking L1 or L2 is enough to activate the appropriate lexical and morphophonological options within that language, with no activation (and therefore no competition) of the unwanted language. This entails that language selection occurs at the conceptual level and is driven by the message (La Heij, 2005). Language-nonselective lexical access models such as in Hartsuiker et al. (2004) described above, Kroll et al. (2006), and Kello et al. (2000), on the other hand, diverge between the locus of selection being at the lemma level, the phonological level or beyond the phonological level. Conversely, Kroll et al. (2006) argue against a fixed locus of language selection in bilingual speech and claim that the suppression of linguistic alternatives during production depends on characteristics of the bilingual speakers and the communicative context. Additional support in favor of non-selective models of language production is the phenomena of the tip-of-the-tong effect (TOT) and cross-linguistic priming.

The TOT is a momentary inability to retrieve words in one of the languages spoken by a bilingual. Kreiner and Degani (2015) observe that both early and late bilinguals exhibited TOT from long- and short-term language exposure, and claim that the phenomenon can be explained under a combination of two apparently contradicting theories: the Frequency Lag Hypothesis (Gollan et al., 2011) and the Dual-Language activation account (Hermans et al., 1998). The Frequency Lag hypothesis is the

difference in frequency effects bilinguals experience both in comparison to monolinguals in their dominant language and between their dominant and non-dominant languages. The hypothesis claims that this lag in frequency comes from the smaller frequency of use of either languages by the bilingual in relation to that of the correspondent monolinguals, making word access in production and comprehension costlier in comparison. Under a non-selective model of production, the language selection entails a higher number of competing forms in the bilingual system and, consequently, decreases as proficiency increases – in fact, Gollan et al. (2011) take L2 proficiency "as a tool that allows a […] manipulation of frequency" (p. 189). The Dual-Language activation account, on the other hand, defends that both languages are activated during production, and these activation levels are proportional to the frequency of use and choice of each language in speech through processes such as inhibitory control (Green, 1998). The difference in activation is the decisive factor of language selection in production.

The Frequency Lag Hypothesis and the Dual-Language activation account converge in that there is competition between language forms from L1 and L2, rendering TOT observations incompatible with a language-selective model. Additionally, Levelt et al. (1999), as outlined above, establish a "rift" between the conceptual/syntactic domain to the phonological/articulatory domain of speech production, based on TOT observations. The momentary inability to activate the pertinent word form in spite of the availability of its concept is indication that there has been both lexical concept and lemma selection, but not word form retrieval. Thusly, the grammatical properties of the word are already activated as evidenced by speakers retrieving the word's gender and number despite not being able to retrieve the word form itself (Viggliocco et al., 1997).

Cross-linguistic structural priming is the activation of a structure in one language increasing the likelihood of that same structure occurring subsequently in the other. This phenomenon is precisely what supports the stratification in the model of bilingual sentence production by Hartsuiker et al. (2004). According to the authors, cross-linguistic priming is evidence of shared representations at the lemma level: the activation of the nodes related to the passive structure during L2 comprehension made them more readily available for retrieval in L1 production, indicating that such nodes

cannot be considered language-specific[3]. It falls out of the scope of this study of identify the exact locus of language selection, but there is enough evidence to support that it takes place after the selection of lexical concept.

As the model of bilingual production proposed by Hartsuiker et al. (2004) is an adjustment of Levelt et al.'s (1999) lexical access network to bilingualism, it is possible to infer that the scope of automaticity in speech production is similar in L1 and L2: modular processes of grammatical and phonological encoding, articulation, and self-monitoring are statistical in nature and largely automatic, whereas concept preparation and lexical concept selection are attention-based – and, consequently, not automatic processes. In this context, L2 proficiency can be understood as the level of automaticity in the grammatical and morphophonological processes of sentence production before language selection.

1.3. *Bilingual shared representations as a function of L2 proficiency*

Although approaches to L2 proficiency may present conflict concerning the aspects, measuring strategies and even the very nature of linguistic knowledge, they converge in the sense that the processes of development of L2 are a function of L2 exposure, leading to frequent or repeated episodes of linguistic processing. This is not surprising, considering that domain-general cognitive mechanisms of learning and categorization are but statistical reconfigurations of knowledge. In Segalowitz and Hustijn's words, automaticity is the "prime psychological construct invoked for understanding frequency effects and how repetition leads to improvement in L2 skill (or any skill for that matter)" (Segalowitz and Hustijn, 2005; p. 371).

Studies in experimental psycholinguistics of bilingualism depend largely on measures of proficiency, especially concerning the interaction between L1 and L2 in the bilingual's mind. In order to investigate the timeline of sharing representations between the L1 and the L2, Bernolet, Hartsuiker, and Pickering (2013) conducted a series of cross-linguistic structural priming experiments with speakers of L1 Dutch and L2 English. The authors observed that priming effects interacted with the subjects' proficiency in the L2: more proficient speakers were more susceptible to priming than less proficient speakers. Their results suggest that L2 learners depart from item- and language-specific representations of L2 syntactic structures and move on to more

[3] We use the term *language-specific* to refer to features belonging to either L1 or L2 individually.

abstract representations as they increasingly experience episodes of L2 processing. The authors also observed that abstraction of representations affects structural generalization across languages, that is, the acquisition of a node available in the L2, but not the L1, makes it available for both languages. In fact, Souza et al. (2014) and Souza and Oliveira (2014) have provided evidence in support of Bernolet et al.'s (2013) claims, by observing that unlicensed L2 English structures, such as the caused movement alternation and the resultatives, were more widely accepted in L1 Brazilian Portuguese by high-proficiency bilinguals than low-proficiency bilinguals and monolinguals.

It is believed that, initially, representations of L2 structures are item- and language-specific in the bilingual's linguistic system, and they eventually abstract into a shared representation available for both languages. While Bernolet et al. (2013) limit the representation for similar structures between L1 and L2, the studies from Souza et al. (2014) and Souza and Oliveira (2014) allow us to extrapolate these claims and state that all representations are available for all languages. This is not to say that unlicensed structures from the L2 become licensed in the L1 due to L2 proficiency. While their levels of acceptance were higher among high-proficiency bilinguals than low-proficiency bilinguals and BP monolinguals, they were still not as widely accepted as licensed structures in the L1. Their acceptance originates from the structure distribution in the L2, not the L1; unless the novel structure becomes a part of the L1 system within the speaking community, these occurrences of these L2 structures in the L1 will tend to stay at mid-acceptance levels.

The results from Bernolet et al. (2013) bring an important consideration for the bilingual production model proposed by Hartsuiker et al. (2004). The model predicts structural priming effects between languages, but Bernolet et al. (2013) point out that this only holds true for high-proficiency bilinguals, whose mental representations are "actually shared" (p. 301). This adjustment to the model is in accordance with Ullman's (2004) declarative/procedural model, in that less proficient L2 speakers tend to store complete syntactic structures in their declarative memory, while more mature L2 speakers rely on rule-based mechanisms for L2 processing.

The experimental results from Bernolet et al. (2013) suggest that low proficiency can be understood as the speaker's higher level of dependence on attentional control to process L2. The originally automatic processes of grammatical and

morphophonological encoding in these bilinguals rely on working memory to a greater extent than do high-proficiency bilinguals, reflecting the fact that the bilingual's linguistic knowledge has not yet made the shift from declarative to procedural memory. Ullman (2004) indicates that the nature of information stored in the declarative memory is mostly (if not all) arbitrary and item-specific, as opposed to the procedural memory, which stores "context-dependent stimulus-response rule-like relations" (Ullman, 2004; p. 237). Representations of L2 structures stored in the declarative memory do not provide the linguistic system with the abstract grammatical features of the language, necessary for the bilingual to start deriving rules that can be generalized over the linguistic system as a whole. The apparent inability of low-proficient bilinguals to generalize constitutes the constraint on distributional learning under investigation in this study: representations must be abstract and general, rather than item- and language-specific, for the bilingual to be able to infer rules from item distributions in the L2.

1.4. *Distributional learning*

The hypothesis of this dissertation is that L2 proficiency constrains bilingual representational sharing in the sense that proficiency is an indication that learned structures from the L2 are no longer item- or language-specific, and their distributional and combinatorial properties can be generalized to the linguistic system as a whole. In brief, we hypothesize that L2 distributional learning can only take place in late stages of second language proficiency. Low-proficiency bilinguals are not able to infer rules and generalizations from L2 structures precisely because they are represented as arbitrary items in declarative memory.

Based on work by Saffran et al. (1996), Aslin and Newport (2012) referred to statistical learning as "the process by which learners acquire information about distributions of elements" after observing that statistical cues alone were sufficient for infants to extract information about statistic coherence of samples from their experimental corpus (p. 171). Although there is robust literature on statistical learning in first language acquisition, this mechanism is not limited to language: Aslin and Newport (2014) report studies that tested infant learnability of distributions of musical tone and images, as well as statistical learning among non-human species. This is evidence that this type of learning is believed to be modality-, domain-, and species-

general. Naturally, there are further and more complex statistical computations as well as social pressures and communicative skills necessary for the development of language, which are unique to human language acquisition (Aslin and Newport, 2014).

Although statistical cues have proven to be sufficient for infants to process linguistic input into its underlying components – that is, to infer generalizations from item distribution in the corpus –, Saffran et al. (1996) do not suggest that they are the exclusive determiner for the processing to occur. In fact, the authors claim that language development takes place from a number of other cues that may be correlated with statistical information. However, statistical learning must present a set of restrictions in order to avoid what is called a computational explosion, which refers to the overwhelming number of statistical computations that can be done from a complex set of input (Aslin and Newport, 2014; p. 90).

An important question concerning statistical learning is what it is that causes a speaker to induce a rule based on sometimes sparse evidence. An explanation can be found in terms of gradients of generalization, which may be based on sensory similarity or on repetition-based rules. A distinction often raised between statistical and rule learning is that the first operates at surface levels, while the latter operates at a deeper level, involving abstract patterns (Marcus et al., 1999). However, this distinction raises yet another issue of what induces one or the other type of learning – which motivated Aslin and Newport (2012) to argue in favor of "a single statistical mechanism with a gradient of generalization" (p. 95), depending on the scope of the generalization allowed by exposure to input. For instance, speakers can infer from syllable transitional probabilities both licensing rules for position in a word and punctual abnormalities in the input (that do not result in generalizable information).

Gerken (2006) argues that every learning task involves a gradient of generalization (p. 94). Word categories, for instance, are believed to be inferred from an inventory of relative positions of words in sentences (note that absolute positions are not suitable for such an inference because position varies in languages: articles precede nouns in English, but the position of the subsequent noun changes according to possible intervening words). Implicit learning accounts also argue that every episode of linguistic processing entails learning for the linguistic system (Bock and Griffin, 2000). In fact, in the Dual-Path model of sentence production proposed by Chang et al. (2006), word classes and syntactic categories are derived from prediction error in

sentence production. In line with the assumptions of this model are the findings from Reeder et al. (2013) that adult speakers can both extrapolate and restrict generalizations from the same corpus (i.e. they correctly abstract grammatical information) based on distributional information. The convergence of frequency-based theories such as implicit and distributional learning supports the hypothesis of this study that high-proficiency adult L2 processing undergoes processes analogous to those of children first language development.

The bulk of studies of distributional learning focuses on infants' or toddlers' first language acquisition using auditory stimuli (Saffran et al., 1996; Aslin et al. 1998; Mattys et al., 1999; Maye et al., 2008; among others), while little has been proposed concerning the role of this mechanism in late bilingual language processing. Similarly, the literature on bilingual linguistic system integration has expanded until the sharing of representations between L1 and L2. Therefore, one of the contributions of this study is to include late bilingualism in frequency-based theories of language learning and processing.

1.5. *Structural priming*

Bock (1986) defined structural priming (also referred to as syntactic persistence or structural priming) as "the tendency to repeatedly employ the same syntactic form across successive utterances" (p. 356). In her seminal paper, the author manipulated priming effects of datives (double object and prepositional object) and transitives (active and passive) on three picture-description tasks disguised as memory tasks. Experiments 1 and 2 presented subjects with a first set of auditory sentences and images to memorize, and a second set of stimuli containing the target experimental items as well as some others from the first set. Subjects were supposed to repeat the sentences and describe the images, and, afterwards, indicate whether the items from the second set had been previously seen. However, the double-set design caused subjects not to pay close attention to (and, therefore, fully process) items that were immediately recognized as new, mitigating possible priming effects. Therefore, experiment 3 was designed as a running recognition task where each prime had a chance of appearing later, forcing subjects to fully process all primes in order to perform well in the cover memory task. In addition to the passive and active primes paired with their counterpart images, Bock (1986) manipulated the position of the agent

(balanced between left and right) in experiment 3, as attempt to elucidate the absence of priming effects for human agent events in experiments 1 and 2. Results show an increase in the number of passives produced as a consequence of the position of both human and non-human agents as well as structural priming effects of the infrequent alternatives (passives and prepositional objects).

Pickering and Ferreira (2008) reported that more than one hundred studies had used structural priming manipulations since Bock (1986), and it is safe to say that this number may have at least doubled in the 10 years between Pickering and Ferreira's publication and this dissertation. Fortunately, the plethora of studies available provide a continuously better understanding of the mechanisms underlying structural priming. Studies such as Bock and Griffin (2000) and Chang et al. (2000) have not only successfully replicated Bock's (1986), but have also showed that structural priming lasts over longer lags, that is, the effects extend over the first target structure produced. Likewise, studies including Pickering and Branigan (1998), Hartsuiker et al. (2008), Bernolet et al. (2013) have found that identity between prime and target verb magnify priming effects. These results are informative not only for understanding structural priming itself, but also because they also contribute to the discussion about accounts of language learning. Section 1.6 below offers a description of the two main competing accounts.

1.6. *Activation-based and implicit learning accounts of language learning*

There are fundamental differences in the way activation-based and implicit learning accounts of language learning explain structural priming. The main point of contention concerns the consequences of structure activation to its distribution (and, consequently, generalizations over language): while activation-based accounts attribute priming to residual activation in the system, implicit learning accounts define the effects as a rule abstraction from the linguistic expression processed.

As described in section 1.2, lexical access accounts of speech production presume a lemma stratum where "lemmas of nouns and verbs are connected to combinatorial nodes specifying the lemmas' subcategorization frames" (Bernolet et al., 2016; p. 99). Therefore, processing of a passive sentence such as (3) involves the activation of the lemma *strike*, but also activates the combinatorial nodes of lexical

category, present tense, progressive aspect, third person, and naturally, the passive and its argument structure.

3. The house is being struck by lightning.

According to activation-based accounts, structural priming occurs because the activation of these nodes decay, but do not disappear immediately. After processing a sentence such as (3), the activation levels for the combinatorial nodes of a transitive verb are higher than for the competing structural nodes, facilitating its selection. This approach to structural priming was first proposed by Pickering and Branigan (1998) as the lexical activation model. Malhotra et al. (2008) proposed a formalization of the account, calling it the trailing-activation model, in which "each episode of training leaves a memory trace based on the units it activates, recorded as a fixed amount of adjustment to the system" (p. 657). Its network architecture (figure 3) consists of two layers, each an independent cognitive module: layer 1 is responsible for syntactic processing (i.e. grammatical constructions), while layer 2 is responsible for lexical processing (i.e. verbs). There are two types of connections in the model: between layers and between nodes in a layer. Connections between layers take place in an intermediate layer consisting of a cognitive module that provides binding nodes: activation-based short-term memory (STM) for associations between layers 1 and 2. While STM establishes mutually excitatory connections between the two layers, connections within nodes in a layer are mutually inhibitory in a winner-take-all (WTA) dynamic: the nodes of a particular layer compete for maximum activation, and the winning node suppresses the other nodes completely.

Both structural priming effects and long-term learning are byproducts of hysteresis[4] in the nodes: activation leaves memory traces in the system, which are recorded by means of STM processes and incremental adjustment to inputs of the winning nodes – analogous to Hebbian learning. This adaptation is also responsible for forgetting processes, which are logical developments of the suppressions resulting from WTA dynamics and necessary processes for any capacity-limited memory system (Malhotra et al., 2008).

[4] Tendency of a system to maintain its properties in the absence of the stimulus that caused them to be.

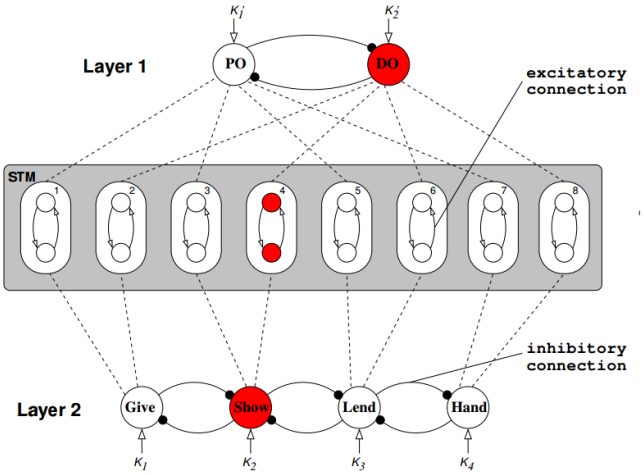

Figure 3 - Trailing-activation network

The trailing-activation model presupposes an independent memory system where linguistic rules are not extracted from each episode of linguistic comprehension, attributing priming effects to “unsupervised, associative learning which leads to traces of activation in the system” (Malhotra et al., 2008; p. 657). It predicts that residual activation (and, consequently, structural priming) is short-lived and stronger for lexically overlapping items since there are trailing activation links from the lexical item to syntactic nodes as well as activation of the syntactic nodes themselves (Pickering and Branigan, 1998). Finally, this model accounts for cumulative priming effects as a result of Hebbian learning: if activation of a given set of lexical nodes favors activation of their binding nodes, these neural pathways will be strengthened over time.

Even though Guimarães (2016) attributed the difference in production of passives between bilinguals and monolinguals to the shared-syntax account of bilingual production (Hartsuiker et al., 2004), this account can only explain the phenomenon to a certain extent. The results reported do add to robust evidence of shared mental representations in bilinguals, but the influences of distributional patterns from the L2 cannot be accounted for in a model of episodic traces on the lemma nodes. Residual activation can hardly account for the difference in the production of passives between bilinguals and monolinguals in a sentence elicitation task conducted solely in the L1 because the task involved no explicit L2 activation. Instead, what was observed was a long-term influence from experience with the L2. Although the trailing-activation model, which supports the bilingual production model in Hartsuiker et al. (2004), could

explain these effects of cumulative priming due to Hebbian learning, constructions that are more frequent would be expected to have greater structural priming effects due to easy of retrieval. Literature shows that this is, in fact, the opposite: more infrequent structures tend to prime more strongly (Bock, 1986; Chang et al., 2006; Jaeger and Snider, 2007; Jaeger and Snider, 2013).

An attractive alternative to activation-based accounts would be to take the difference in production observed in Guimarães (2016) as a suggestion that language learning and structural priming share the same mechanisms. In fact, implicit learning accounts of language learning propose that the speaker adjusts his or her production system as a function of experience with episodes of linguistic comprehension (Bock and Griffin, 2000). Chang, Dell, and Bock (2006) proposed a dual-path model of sentence production that accounts for linguistic productivity, and thus for adult language learning and structural priming. The model eliminates the need for any innate abstract syntactic knowledge (cf. Chomsky, 1959) to explain productivity, otherwise arguing that syntactic rules are abstracted from sequences of words.

Syntactic abstractions in implicit learning accounts come from adjusting predictions about what the speakers hear. The system adjusts prediction weights based on the difference between its predicted output and the correct output via backpropagation, i.e. the adjustment in the weights of hidden units in a network so that the model learns arbitrary pairings of input and output. A type of learning through backpropagation is the simple recurring network (SRN), which is a "[…] a feed-forward three-layered network (input-to-hidden-to-output) [that] also contains a layer of units called the context that carries the previous sequential step's hidden-unit activations" (Chang et al., 2006; p. 234). Chang et al. (2006) consider SRN an important part of theories of language learning because it sequentially accepts inputs and predicts outputs, thus using information about both past and present to predict the future, as well as providing good accounts for generalizations. The SRN for incremental word prediction works by taking comprehended words (*cwords*) as input, predicting the next word (the output) and comparing it to the next heard word – this word then becomes the input for the next prediction cycle. The system is only able to depart from simple word prediction to sentence production because the process involves message representations – concepts represented in an event-semantics frame.

The dual-path model is so named because it entails two separate pathways for both prediction and production: the sequencing system and the meaning system (figure 4). The sequencing system is designed to learn information so that it produces words in a syntactically acceptable way, so a byproduct of its processes is creating lexical and syntactic categories from comprehension. Knowledge gained from comprehension transfers readily for production when there is no external input (which has consequences for accounts of self-monitoring, for example). The meaning system contains the message – concepts, event roles, and their bindings (temporarily increasing weights between concepts and roles). Although there is some debate about whether event roles are pre-defined or arise from properties of concepts (McRae, 1997), roles in the dual-path model are given by event-semantics and assigned to concepts through activation – roles that are more prominent are assigned to concepts that are more prominent. The sequencer can only access event roles in the meaning system; the event role is linked to the concept, allowing the words to be correctly sequenced.

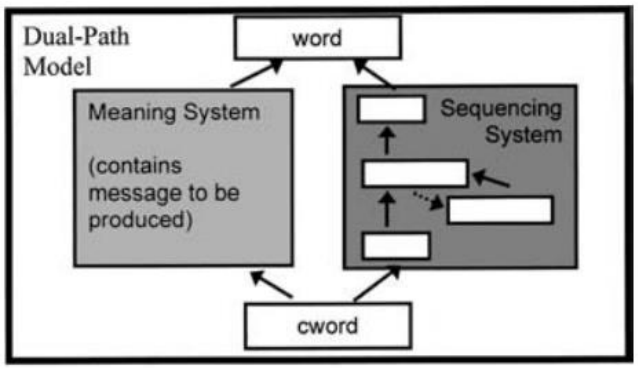

Figure 4 - Dual-path model

A consequence of the structure of the model for the production of passives is that the transitive event is comprehended as a whole, and the most prominent concept (i.e. the one with higher activation levels) will take the most prominent event role (i.e. the one first mentioned), thus defining the structure of the sentence. For instance, the event shown in figure 5 is expected to be expressed with an active or a passive

structure depending on whether the police officer or the thief is more prominent to the speaker[5].

Figure 5 - Picture used to describe the event "arrest"

This is in line with the results reported by Gleitman, January, Nappa, and Trueswell (2007). In their attention manipulation study, they observed that speakers tended to produce the cued participant first, regardless of the fact that it would entail use of the least preferred structure. Their analysis of reaction times also showed that the event is completely represented before it is uttered, in accordance with Griffin and Bock's (2000) interpretation:

> "The evidence that apprehension preceded formulation, seen in both event comprehension times and the dependency of grammatical role assignments on the conceptual features of major event elements, argues that a wholistic process of conceptualization set the stage for the creation of a to-be-spoken sentence" (Griffin and Bock, 2000; p. 279).

The dual-path model makes some important assumptions for bilingual representational sharing. First, it assumes that processing is learning, since the mechanisms employed in early language acquisition functions throughout adult life. Second, it assumes that learning occurs when a predicted word deviates from a target word, both production and comprehension – learning takes place through prediction error.

1.7. *Surprisal and cumulativity*

Surprisal and cumulativity are two key aspects of priming that unfold from the perspective that the phenomenon is a result of implicit learning. As previously

[5] In fact, the model assumes that event role assignment processes are analogous to recognition of objects in the visual space (Chang et al., 2006; p. 237).

mentioned, this account of language learning assumes that every episode of linguistic processing entails an update of the distribution of the structure, and this amount of learning determines the probability of using it afterwards (Jaeger and Snider, 2007, Bernolet et al., 2016).

It is argued that this syntactic persistence supports learning through linguistic processing, as stated by error-based accounts, rather than activation-based approaches in which recency of activation is the cause of structural repetition (Bock and Griffin, 2000). Jaeger and Snider (2007) further divide syntactic persistence into surprisal-sensitivity and cumulativity. Surprisal is defined as the inverse-frequency effect, which predicts that less frequent constructions cause higher prediction error and, consequently, prime more strongly. Cumulativity refers to how many prime structures have been processed in a conversation until it is used by the speaker, thus predicting that the more instances of a construction a speaker has produced or comprehended, the more likely they are to produce that structure in the future. Surprisal and cumulativity may sound contradictory at first, since the first predicts stronger priming effects for less frequent constructions and the latter predicts more likelihood of production after more instances of processing. However, it is important to notice that cumulativity in conversation does not mean the construction loses its low frequency status; consistent use of passives in a conversation, for instance, does not rearrange the probability distributions of the passives in relation to actives.

Error-based accounts explain structural priming effects in terms of learning (rather than transient activation), given that change in performance persists over longer lags and generalizes to new utterances with different words. In Bock and Griffin's (2000) words,

> "[…] the relevant kind of learning appears to be implicit or procedural, inasmuch as it does not depend on specific intentions to replicate a sentence's structures in new words, does not require an effort to remember the priming sentences (Bock, 1986), and does not require explicit attention to the form of a priming sentence" (p. 187).

This study relies on structural priming inasmuch as it has been shown to reflect cross-linguistic influences, such as in Hartsuiker (2004). Therefore, we assume cross-linguistic structural priming effects to analyze surprisal effects on different profiles of speakers (monolinguals and bilinguals with low and high L2 proficiency).

1.8. *Syntax of oral production*

Jaeger and Snider (2007) emphasize the importance of naturalistic data to psycholinguistic studies, as observations of natural language production cannot be considered an artifact of task-related learning. Besides circumventing laboratorial limitations, the use of spontaneous oral data contributes to the study of language as an emergent phenomenon directly reflecting underlying cognitive processes – as opposed to the analysis of well-formed written sentences that putatively depict competent linguistic knowledge.

It is widely accepted that speech is the natural modality of language, whereas writing constitutes a technological feat. However, the vast majority of linguistic analyses have used written language as the object of research and generalized the findings over to oral production, in spite of the fundamental differences between the two modalities. This overgeneralization may find support in the chomskyan theory that well-formed sentences are a much better reflection of a speaker's linguistic competence than oral performance, which is subject to slips of the tongue and misinterpretations (Chomsky, 1965). The discrepancy between the syntactic well-formedness observed in writing, but not in speech, is explained away with traces, i.e. null copies of the linguistic elements that would eventually be missing in oral production.

Analysis of spontaneous speech is fundamentally incompatible with this type of approach, since the empirical nature of such studies does not allow research to rely on abstract categories that, as expected, would not show on any oral corpora. Once traces or subjacent structures are not eliminated from analysis, we are faced with the fact that the fundamental differences between written and spoken language require different processes of linguistic analysis: speaking and writing differ in terms of time realization, possibility of immediate feedback and meaning negotiation, physical availability, lexical density, structure complexity, and so forth (Raso, 2013). The most prominent of such differences may seem obvious at first but has important theoretical implications: there needs to be an acoustic signal for linguistic expression to be considered speech.

Speech is based on a number of extra-linguistic factors that can only take place in a situation of interaction (e.g. gestures, shared knowledge and context, physical location, facial expressions) as well as on the linguistic expression itself to convey meaning; any attempts to transfer the linguistic expression alone to written language

would result in an unsuccessful production. In the absence of prosody, a structure such as (1) would be misinterpreted:

4. ***não tem que colocar uns espelhos aqui***
   no have to put some mirrors here
   '[You] don't have to put any mirrors here.'

In writing, the scope of negation (*não*) would be over the subsequent proposition, indicating that someone should not put up any mirrors in that place. However, prosodic analysis shows a non-terminal break between "não" and the rest of the sentence, shifting the scope of negation to the previous rather than the following proposition:

5. ***não / tem que colocar uns espelhos aqui //***
   no / have to put some mirrors here //
   'No, [you] have to put some mirrors here.'

Prosody then provides crucial information about the meaning of the proposition: it changes to a rejection of what had been said before, showing the speaker's position in favor of putting up the mirrors. Indeed, Raso and Mello (2014) argue that prosody is the main vehicle for linguistic functions and it should not be viewed solely as paralinguistic information. The acoustic signal, the most fundamental difference between oral and written production, is precisely what makes it possible for prosody to play its linguistic role in language production.

Raso and Mello (2014) defend that the sentence cannot serve as the reference unit for speech syntactic analysis, given the number of para- and extra-linguistic factors and the fundamental differences between oral and written production (as discussed above). Although there is still some debate over what the most suitable reference unit for speech is, utterances and information units are a reliable compass for syntactic analysis of oral production. It is important to highlight that prosody is what determines the terminal or non-terminal character of the breaks that divide utterances into information units; therefore, there can be no oral production analysis without the sound itself.

Oral corpora analysis presents itself as an invaluable resource to study the passive construction, given the role pragmatics plays in the analyses of both. Under the view of Construction Grammar (Goldberg, 1995), passives are complex structures

that stem not from expression of basic human experience, but from a pragmatic need to rearrange the informational structure of the linguistic expression. Oral cues – namely, prosody – can potentially be a source for further elucidating the conditions that constrain the occurrence of the passive construction, and how they may vary in L1 and L2 production. Moreover, it adds to the perspective of language as modality instead of a unified and homogenous system. The results from Guimarães (2016) in section 1.1 clearly illustrate the need for such an approach: significant differences were observed in production of passives by monolinguals in the written and the spoken tasks, evidence of the different aspects and (extra) linguistic restrictions that affect each of the language modalities.

2. METHODOLOGY

The hypothesis that late L1 BP L2 English bilinguals share distributional learning mechanisms between L1 and L2 departs from a number of assumptions that must be properly documented before any conclusions can be made about the underlying mechanism of language learning and process in bilinguals. These assumptions include the status of the passive structure in BP, its properties of structural priming (surprisal-sensitivity, cumulativity, and lexical boost), L2 proficiency measures, and different levels of representational sharing between low- and high-proficiency bilinguals. Only after properly investigating these assumptions will we be able to confirm or reject the main hypothesis of this dissertation.

While most of the literature about surprisal-sensitivity and cumulativity involves English in both within- and between-language structural priming, it is only recently that it has been studied in BP under implicit learning accounts. Some of the recent studies include Teixeira (2016), who observed effects of structural priming of passives on the production of children, but not of adults; Belavina Kuerten et al. (2016), who found priming effects on the production of dyslexic children; and Kramer (2017), who observed a decrease in the reading times of the passive after reading another passive among elementary school children.

All three studies analyzed experimental data, which controls the conditions for the structural priming effects to take place. However, a complete understanding of structural priming in BP must include analyses of naturalistic data which, according to Jaeger and Snider (2007), “cannot be an artifact of unnatural distributions that may cause explicit learning rather than implicit, highly automatic learning” (Jaeger and Snider, 2007; p. 28). Another implication of using data from spontaneous speech is the adoption a theory of oral syntax that rejects the existence of traces (of verbal complements, for example) or subjacent structures preceding oral production. Unlike the studies of priming in BP aforementioned, this dissertation analyzes the passive as an independent construction that reflects the speaker’s perspective of the event comprehended (Raso and Mello, 2012; Goldberg, 1995; Ellis, 2003; Griffin and Bock, 2000).

As an attempt to contribute to the formation of literature on the analysis of structural priming in BP in naturalistic data, the first study is an examination of the oral corpus of BP C-Oral-Brasil I. Following the analyses in the voice alternation

experiments reported by Jaeger and Snider (2007), study 1 investigated the properties of surprisal-sensitivity and cumulativity in structural priming effects of the passive in BP.

This research also presents an experimental component, designed to contrast monolinguals, low-proficiency and high-proficiency bilinguals. The distinction between the subjects has two main motivations. First, since the passive is significantly less productive in BP than in English, the inverse-frequency effect leads us to believe that monolinguals are more sensitive to the priming effects of the structure than bilinguals. Second, the results from Bernolet et al. (2013) suggest that low-proficiency bilinguals' surprisal-sensitivity may differ from that of high-proficiency bilinguals, depending on the extent of representational sharing – which, as discussed, is motivated by experience with the L2.

We expect that the corpus analysis and the behavioral experiment will provide answers to the assumptions underlying the main hypothesis. Section 2.1 offers a summary of findings from Jaeger and Snider (2007), followed by the complete report of studies 1 and 2. Afterwards, section 2.5.1 brings an overview of the studies by Bock (1986) and Guimarães (2016), which guided the experiment in study 3, described in section 2.5.

2.1. *Jaeger and Snider (2007)*

In an attempt to contrast transient activation and implicit learning accounts of syntactic persistence, Jaeger and Snider (2007) conducted two studies on surprisal and cumulativity of voice alternation using the voice alternation data set from the Penn Treebank portion of the Switchboard corpus (Marcus et al., 1999). Transient activation accounts of syntactic persistence predict short-lived priming effects and significant influence of prime-target verb identity due to the activation of lemma representations of both the passive construction and the verb, while implicit learning accounts predict slower decay of priming and little influence of verb identity – the effects originate from recalibration of distributions of the constructions over episodes of linguistic processing (both comprehension and production). Thus, Jaeger and Snider (2007) relied on the properties of surprisal-sensitivity and cumulativity to distinguish between transient activation and implicit learning accounts of syntactic persistence.

The first study focused on the effects of surprisal-sensitivity in voice alternation. Surprisal-sensitivity is defined as the log inverse probability of occurrence of the structure in a given context, supported by the observation that more infrequent structures tend to prime more strongly (cf. Bock, 1986). Independent variables considered in structure choice included prime and target verb bias (i.e. the conditional probability of the passive occurring, given the verb), distance between prime and target (to control decay); verb identity between prime and target was used as a control factor; finally, speakers were added as a random factor to control for individual variability in the production of passives.

Jaeger and Snider (2007) found significant effects of all factors: prime verb bias and distance between prime and target yielded negative coefficients, while target verb bias and verb identity yielded positive coefficients, as shown on table 1. The observation that prime verb passive bias shows a negative correlation with the production of passive targets supports the prediction that syntactic persistence is surprisal-sensitive (figure 6).

Predictor	*Parameter estimates*			*Wald's test*		$\Delta_x(\Lambda)$*-test*	
(independent variable)	Log-odds	S.E.	Odds	Z	P	χ^2	P
prime V = target V	1.754	0.259	5.75	6.8	$\ll 0.001$	46.8	$\ll 0.001$
target verb pass. bias	1.212	0.025	3.35	47.9	$\ll 0.001$	3489.7	$\ll 0.001$
prime-target distance	-0.060	0.026	0.94	-2.3	< 0.04	6.7	< 0.01
prime verb pass. bias	-1.205	0.404	0.30	-3.0	< 0.002	9.2	< 0.003

Table 1 – Summary of passive surprisal analysis (Jaeger and Snider, 2007)

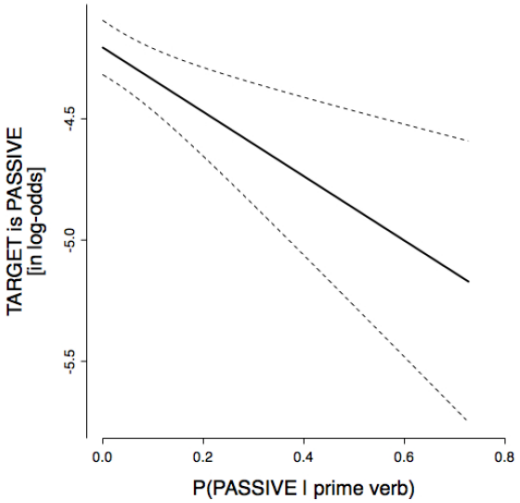

Figure 6 - Prime surprisal based on prime verb's passive bias (Jaeger and Snider, 2007)

The second study addressed the issue of whether syntactic persistence is cumulative, that is, the more prime passive structures are comprehended and produced in the conversation, the more likely it is to be produced later. In addition to the same variables used in the surprisal-sensitivity study, within- and between-speaker cumulativity of active and passive structures were added as independent variables. Jaeger and Snider (2007) predicted that the number of actives processed (comprehended and produced) would decrease the number of passives produced, and vice-versa.

There was a significant effect of both passives produced within- and between speaker, and only a small effect of actives produced by the same speaker (Table 2). The difference in effect of passives versus actives produced by the same speaker is in accordance with the surprisal-sensitivity hypothesis that infrequent structures prime more strongly.

Predictor	*Parameter estimates*			*Wald's test*		$\Delta_x(\Lambda)$*-test*	
(independent variable)	Log-odds	S.E.	Odds	Z	P	χ^2	P
prime=*passive*	1.036	0.169	2.82	6.1	$\ll 0.001$	35.1	$\ll 0.001$
prime V = target V * prime=*pass.*	1.757	0.259	5.75	6.8	$\ll 0.001$	46.8	$\ll 0.001$
target verb pass. bias	1.212	0.025	3.35	47.9	$\ll 0.001$	3490.1	$\ll 0.001$
prime verb pass. bias * prime=*pass.*	-1.205	0.404	0.30	-3.0	< 0.002	9.2	< 0.003
prime-target distance * prime=*pass.*	-0.059	0.026	0.94	-2.3	< 0.04	6.7	< 0.01
prec. passives within-speaker	0.107	0.016	1.12	6.6	$\ll 0.001$	36.9	$\ll 0.001$
prec. passives across-speaker	0.055	0.020	1.05	2.8	< 0.02	7.6	< 0.006
prec. actives within-speaker	-0.008	0.002	0.99	-4.1	$\ll 0.001$	16.2	$\ll 0.001$
prec. actives across-speaker	-0.001	0.002	0.99	-0.6	< 0.6	0.3	> 0.5

Table 2 - Summary of passive cumulativity analysis (Jaeger and Snider, 2007)

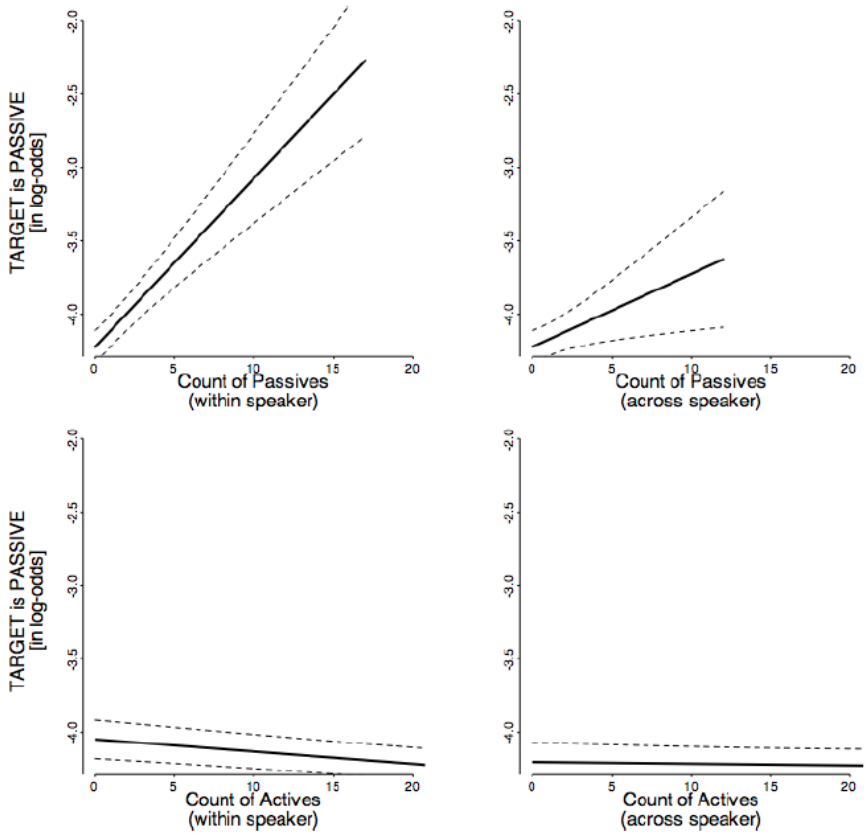

Figure 7 - Cumulativity in passives (Jaeger and Snider, 2007)

2.2. *Study 1: surprisal-sensitivity of passives in BP*

Evidence from corpora and experimental studies has shown that the passive structure is significantly less productive in BP than in English (Guimarães and Souza, 2016; Guimarães, 2016; Duarte, 1990). Thus, it is reasonable to expect that surprisal estimates, surprisal-sensitivity and cumulativity effects of the passive structure in BP will differ greatly from the data in Jaeger and Snider (2007). Study 1 conducted the same surprisal-sensitivity and cumulativity analysis from Jaeger and Snider (2007) on C-Oral-Brasil I, a corpus representative of the diatopic variation of the Brazilian Portuguese spoken in the state of Minas Gerais, in Brazil. C-Oral-Brasil I is so far composed of 263,000 words in 139 texts of informal speech equally divided into monologues, dialogues, and conversations. Because one of the predictor variables is between-speaker cumulativity, only the dialogues and the conversations have been analyzed.

2.2.1. Data

A total of 12418 verbs were extracted from the dialogues and conversations in the C-Oral-Brasil I. First, verbs that do not participate on the voice alternation and verbs that are fixed expressions (e.g. *be supposed to*, *be born*) were eliminated. Verbs with under 10 occurrences in the entire corpus were also excluded, to ensure surprisal calculations were reliable. After, passives and actives were classified as such if they presented the copula verb *ser* (be) followed by a verb in the participle form, or were a transitive verb followed by a direct complement NP, respectively (verbs with oblique complements in BP do not allow passivization). The classification yielded a total of 12307 actives and 111 passives (0.009% of the structures). Sentences 6 and 7 are examples of the passive and active alternation for the verb *construir* (build), extracted from the corpus:

6. / ***é construído toda uma agenda lá na região*** /
   / is built all an agenda there in the area /
   'The agenda is set in the area'
7. / ***cada um foi construir sua casa*** /
   / each one went build their house /
   'Everyone built their own house'

2.2.2. Method

Verbs were extracted from the corpus using the gsubfn package (Grothendieck, 2018) in R (R Core Team, 2013), following the procedures outlined by Gries (2009). The analysis included only verbs with passive primes, as the surprisal-sensitivity from actives (and other preferred structures, such as double-object datives) are known not to show priming effects, that is, the production of actives is not increased by the processing of an active prime (Gleitman et al., 2007; Jaeger and Snider, 2007; among others).

The independent variables from the first study on voice alternation in Jaeger and Snider (2007) were maintained: passive biases of prime and target verbs, lexical identity between prime and target; and distance between prime and target. Passive biases were calculated based on the conditional probability of the occurrence of a passive given the verb[6], and the unit of distance in this study is given in constructions

[6] Verbs and biases are available in Appendix 1.

(cf. Goldberg, 1995), with each finite passivizable verb constituting a distance of 1. Speakers were considered random variables, to account for individual rates of passive production and the lack of extralinguistic information normally controlled for in psycholinguistic analyses. The dependent variable was choice of structure in the target sentence, active or passive.

Following the well-document effect of less frequent expressions priming more strongly (Bock, 1986; Chang et al., 2006; Jaeger and Snider, 2007; Jaeger and Snider, 2013), the distinction in passive distributions in BP and English (Guimarães and Souza, 2016) lead us to predict that priming effects of the passive in BP will be stronger than in English. Prime verbs with smaller biases are expected to make the passive more likely in the target. The lexical identity control may either not have effects at all or increase the likelihood of passive targets due to strength of lexical item activation (cf. lexical access accounts) or explicit memory (cf. implicit learning accounts). Target verb bias is also expected to have a positive correlation with the choice of passives, following their larger number of occurrences in the passive throughout the corpus.

2.2.3. Analysis

The data was analyzed using mixed logistic regression models in R (R Core Team, 2013), using the lmer and lmerTest packages (Bates et al., 2015; Kuznetsova et al., 2017), which accommodate the random variable (speakers) and the repeated measures nature of corpus analysis. For each target structure in the passive, each of the independent variables was analyzed: passive bias, identity with and distance from the prime verb, and target verb bias.

Although prime verb bias showed a negative log-odds coefficient (-0.9400), it did not have a significant effect on structure choice ($Z = -0.526$, $p = 0.599$). There were significant effects of target verb bias ($Z = -9.318$, $p < 0.0001$) and lexical identity ($Z = -7.674$, $p < 0.0001$); however, their log-odds coefficients had opposite effects from what had been predicted: negative coefficients of -16.6850 for target verb bias and of -2.4845 for lexical identity. Distance was also a significant factor with inverted log-odds coefficient: $Z = 2.232$, $p = 0.0256$, and log-odds of 0.011625. Effects of each predictor are shown in figure 8.

Predictor	*Estimate*	*S.E.*	*Z*	*P*
prime bias	-0.9400	1.7870	-0.526	0.599
target bias	-16.6850	1.7905	-9.318	<2e-16
lexical identity	-2.4845	0.3237	-7.674	1.67e-14
distance	0.011625	0.005209	2.232	0.0256

Table 3 - Summary of passive surprisal in C-Oral-Brasil I

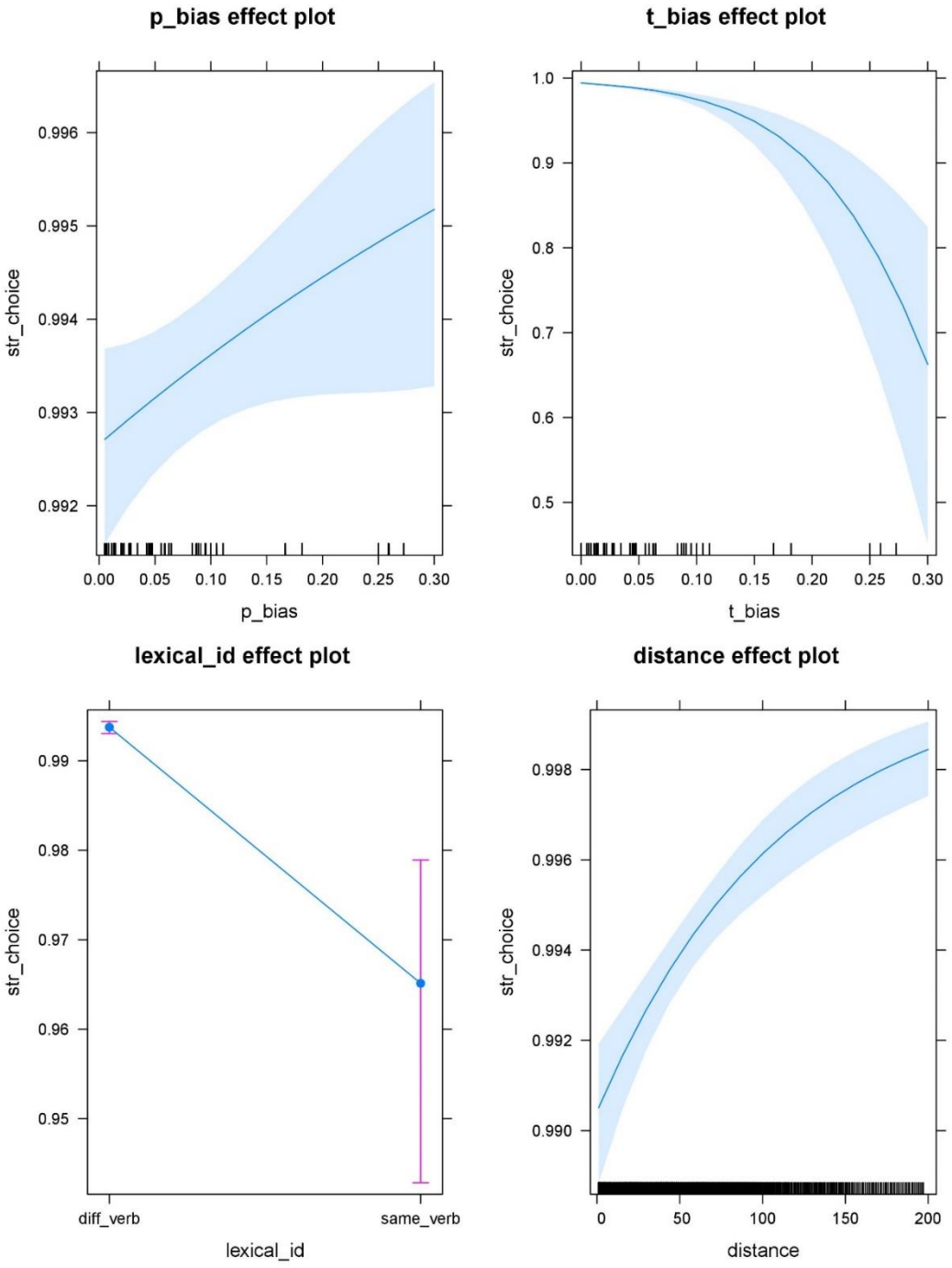

Figure 8 - Individual correlations on structure choice (passive surprisal)

2.2.4. Results

The results from C-Oral-Brasil I differ almost entirely from those observed in the surprisal-sensitivity study by Jaeger and Snider (2007), and none of the predictor variables behaved in the expected way. First, the correlation between prime verb bias and choice of structure was not significant, despite negative. Second, lexical identity had a significant but negative effect on choice structure, which goes against the lexical boost effect observed in many studies (Pickering and Branigan, 1998; Hartsuiker et al., 2008; Bernolet et al., 2013; among others). Third, distance showed a positive significant effect on choice of structure and, by looking at the results prima facie, it could be concluded that speakers tend to produce more passives as the distance between the last heard passive structure increases. Finally, target verb bias showed a negative significant effect on choice of structure, which could be interpreted as the speakers' preference for the passive structure increasing for verbs that occur in that structure more infrequently.

At first glance, these results could be taken as opposing evidence to either of the syntactic persistence models being analyzed. If taken as they are, the data on table 3 suggests that the negative effects of lexical identity could argue against the trailing activation model's claim of ease of access to representation due to activation of both the passive lemma node and the lexical item. Likewise, positive effects of distance between prime and target would lead us to assume that syntactic persistence increases rather than decays over time, while the non-significant effects of prime verb bias would suggest that BP speakers do not show surprisal-sensitivity to infrequent structures.

However, the fact that neither the controls of lexical identity and prime target distance nor the predictor variables of prime and target verb bias showed consistent behavior does not allow us to assume that priming effects took place in the conversations analyzed. Concluding from this data alone that passives do not have priming effects in BP would go against an expressive body of literature that has attested the effect in English corpora (Gries, 2005; Jaeger and Snider, 2007) and cross-linguistically with languages such as Spanish, German, and Dutch: Hartsuiker et al. (2004) found effects of cross-linguistic structural priming from L2 English to L1 Spanish in the production of passives by bilinguals; Melinger and Dobel (2005) found priming effects of previously shown words on the dative alternation in German and

between L1 German and L2 Dutch; Bernolet et al. (2013) found priming effects of L2 English on L1 Dutch genitives contingent to L2 proficiency.

Assuming priming effects failed to take place, the distance factor loses its explanatory power. Instead of being an indication of the rate of decay of the passive prime, it only indicates the number of intervening constructions between two unrelated instances of passives. The positive correlation makes sense, considering that an extremely small number of occurrences of passives in the corpus in general is expected to be dispersed throughout conversations. Although the left-skewed distribution of the distance values in passives might suggest an effect contingent to prime verb bias (an increase in distance would result in a decrease of prime-surprisal of verbs with bigger biases), this was not the case. An analysis with interaction between prime verb bias and distance proved negatively correlated, but non-significant (-0.038643 log-odds, Z = -0.558, p = 0.5766).

A possible explanation for these uninformative results could be the design of C-Oral-Brasil I. The corpus is comprised of monologues, dialogues, and conversations, of which only the last two were included in the study due to the speaker identity control in study 2. However, priming effects in conversations might have suffered from the multiple-speaker configuration, given the impossibility to track each speaker's attention or presence throughout the conversations. It is possible to conjecture that not all of the speakers focused on all the other speakers during the entire time, as often happens in spontaneous conversations; in addition, in some of the conversations one or two participants only speak in the very end of the recording, suggesting that some of the speakers may have been exposed to only a subset of the (already few) instances of passive structure. The Penn Treebank corpus, on the other hand, is comprised of telephone conversations that take place between only two participants, whose focus is directed at one another the entire time. The difference between the setting of the conversations could have been responsible for such discrepant results.

Another possibility may be the properties of the passive structure in Brazilian Portuguese. Similarly to the discrepancy in the distribution of the passive structure between English and BP reported by Guimarães and Souza (2016), a chi squared test revealed that the frequency of passives in both corpora was significant $\chi 2$ (1) = 508.969, $p < 0.0001$:

	C-Oral-Brasil I	*Penn Treebank*	*total*
actives	12307	29007	41314
passives	111	1791	1902
total	12418	30798	43216

Table 4 - Frequencies of passives in C-Oral-Brasil I and Penn Treebank corpora

Guimarães (2016) also reported a significantly smaller number of passives produced by BP monolinguals in the sentence elicitation task. The difference is significant both in comparison to bilinguals' production and to the monolinguals' production in the written sentence elicitation task. The behavior of BP speakers concerning the passive construction differs from what is reported in the literature (Hartsuiker et al., 2004; Dussias, 2003), and qualitative analysis is needed in order to identify the restrictions that might mask structural priming of passives in naturalistic settings.

Additionally, the distribution of passive verb bias in BP and English may account for the sparsity of passive structures in BP and consequent absence of priming effects in relation to English. Since the list of verb biases is not available in Jager and Snider (2007), we calculated the passives biases of transitive verbs in English from the Santa Barbara Corpus of Spoken American English, or SBCSAE (Du-Bois et al., 2000-2005). The SBCSAE is comprised of approximately 249,000 words and was compiled from face-to-face informal interactions between speakers of various locations from the United States. The entire list of verb biases is available in Appendix 2.

Although the number of passive-biased verbs was not expressive in either language (only the verb *prender* (arrest) in BP and the verbs *involve* and *name* in English presented passive biases greater than 0.5), the frequency of entirely active-biased verbs in BP motivated the hypothesis that lexical properties of verbs in BP might explain the scarcity of passives in BP, especially in comparison to English. In fact, 72.5% of the BP verbs show a null passive bias, meaning that they do not occur in the passive at all in the entire data set; in English, approximately 40,2% of verbs are entirely active biased.

The hypothesis that passive verb bias differ significantly between English and BP was examined through a Wilcox signed-rank test for non normal distributions, which yielded significant results ($W = 12872$, $p < 0.0001$).

Corpus	*Min.*	*1st Qu.*	*Median*	*3rd Qu.*	*Max.*
C-Oral-Brasil I	0.00000	0.00000	0.00000	0.00838	0.53846
SBCSAE	0.00000	0.00000	0.08047	0.10000	0.78947

Table 5 - Statistical description of verb biases in BP and English

Given the small percentage of verbs that occur in the passive structure in BP, a possible explanation is that the construction in BP constrains a specific semantic-pragmatic class of verbs. This is not to be confused with processes of grammaticalization of passive expressions, instantiated in English by expressions such as *be born* and *be supposed to.* These expressions cannot be considered passive alternations due to the impossibility to occur in the active alternation whithout changes to meaning or to receive an agentive by-phrase, in spite of the presenting the morphosyntactic configuration of a passive:

8. Giovana will be born in September.
   ? Leda will bear Giovana in September.
   ? Giovana will be born by Leda in September.

9. Iara is supposed to wake up at 6 o'clock.
   ? Taís supposes Iara to wake up at 6 o'clock.
   ? Iara is supposed to wake up by Taís at 6 o'clock.

However, Ciríaco (2011) observed that the passive in BP is an independent sentence pattern, associated to the meaning of directed eventuality and constrained by the event's conceptualization of agency. This means that the occurrence of the structure does not depend on verb class or even lexical item, being fully licensed for agentive events.

It is important to stress that this is a tentative explanation for the distributional difference that is believed to have motivated the absence of priming effects from such a marked structure as the passive in BP. However, the conjecture that the verb selection in the passive structure is semantically or pragmatically motivated would be in conflict with priming effects in BP documented in literature (e.g. Teixeira, 2016).

2.3. *Study 2: cumulativity*

2.3.1. Data

The same data set from study 1 was analyzed for cumulativity effects of the passive structure in the context of a conversation (group or dialogue). As in Jaeger and Snider (2007), cumulativity was calculated from the number of actives and passives processed by the speaker until the production of the target structure.

2.3.2. Method

The cumulativity effect was measured from the number of structures processed preceding the target passive structure in order to assess whether speakers' choice of structure would correlate with recent linguistic episodes in the context of a conversation. Within-speaker persistence was coded as actives and passives produced, and between-speaker persistence was coded as actives and passives comprehended. As in Jaeger and Snider (2007), the predictions were that actives processed would not yield effects on structure choice due to their high frequency (all verbs except for *prender* (arrest) are active-biased in the corpus), but passives processed would increase the likelihood of passive production.

2.3.3. Analysis

As in study 1, the data was analyzed using mixed logistic regression models with speakers as a random variable, which is especially relevant in an analysis taking individual variation as an independent variable. Target verb bias had the same negative significant effects as in the surprisal-sensitivity study ($Z = -11.51$, $p < 0.0001$), and the effect of actives fell within the expected non-significant positive correlation ($Z = 0.751$, $p = 0.453$ for between-speaker persistence and $Z = 1.168$, $p = 0.243$ for within-speaker persistence). Also in line with the results of study 1, but opposite to the predictions, were the effects of passive cumulativity: both showed a negative non-significant correlation ($Z = -1.657$, $p = 0.0975$ for between-speaker persistence; $Z = -1.486$, $p = 0.137$ for within-speaker persistence). Results are summarized in table 6 and plotted in figure 9.

Predictor	*Estimate*	*S.E.*	*Z*	*P*
target verb bias	-18.9995	1.6505	-11.51	<2e-16
actives comprehended	0.001938	0.002580	0.751	0.453
actives produced	0.004183	0.003581	1.168	0.243
passives comprehended	-0.1916	0.1156	-1.657	0.0975
passives produced	-0.2402	0.1616	-1.486	0.137
target V bias * pass. comp.	-2.7325	1.588	-1.721	0.0853
target V bias * pass. prod.	4.2545	1.9734	2.156	0.0311

Table 6 - Summary of passive cumulativity in C-Oral-Brasil I

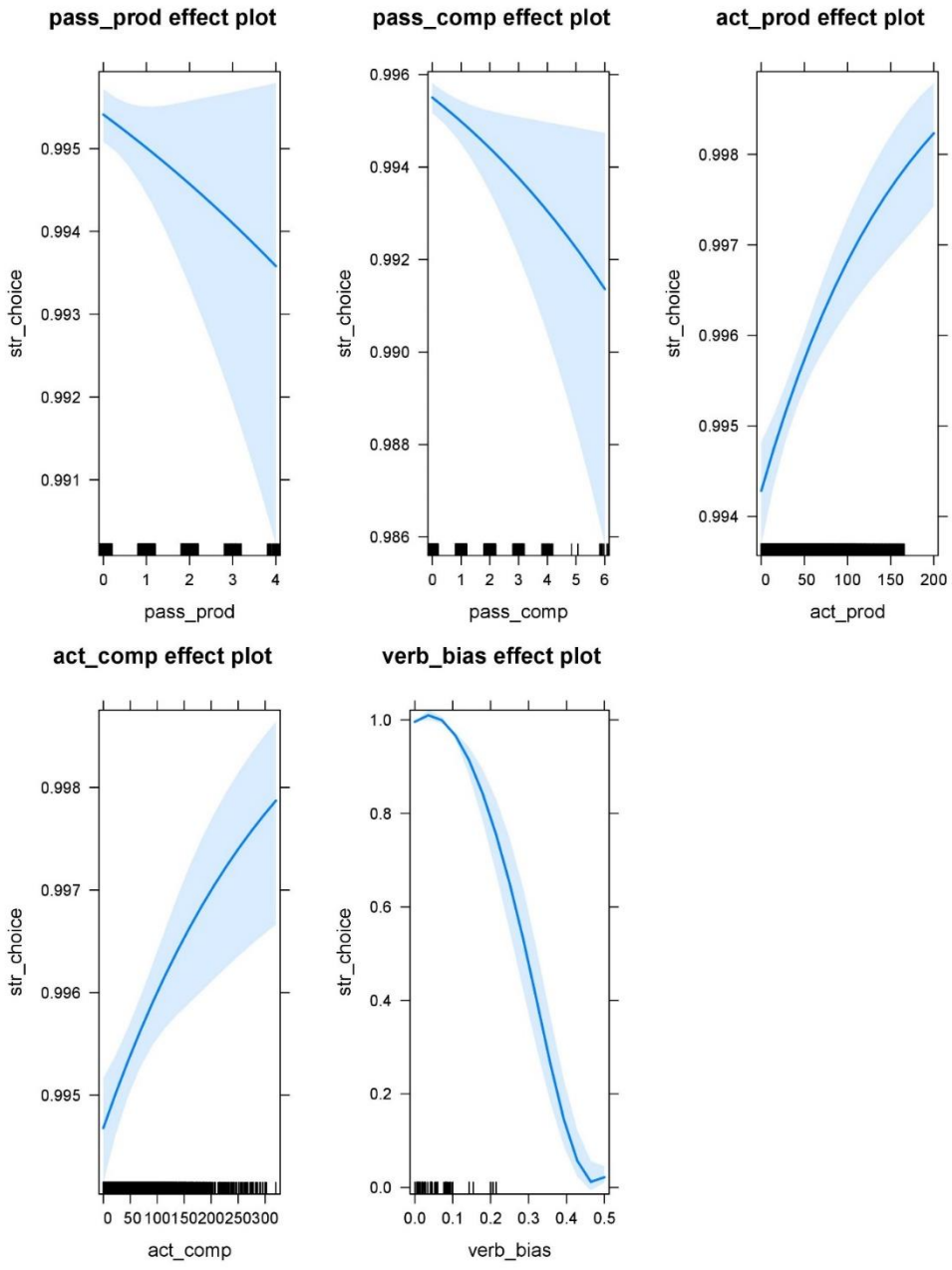

Figure 9 - Individual correlations on structure choice (passive cumulativity)

2.3.4. Results

In light of the results from study 1, caution should be exercised in the interpretation of these results. We depart from the observation that priming effects appear not to have taken place in the data, inasmuch as none of the control variables from the first analysis behaved as predicted. Once again, the negative correlation between passive cumulativity and structure choice can not immediately be taken to mean that the likelihood of a target verb being in the passive structure decreases by every passive structure processed. Instead, target verb bias and passive cumulativity should be analyzed in light of inner biases of passive production – not depending on the speakers themselves, as they are random factors, but on the distribution of the passive construction given the verb passive bias from the entire data set.

From this perspective, the interaction results reported on the lower section of table 6 are somewhat expected. The absence of statistical significance in the interaction between target verb bias and number of passives comprehended (Z = -1.721, p = 0.0853) is in line with the also non-significant prime verb bias effect from study 1: there is no evidence of priming effects having occurred in the data. Conversely, the significant effect observed in the positive correlation between target verb bias and passives produced (Z = 2,156, p = 0.0311) suggest that the effects of within-subject cumulativity increase the likelihood of producing passives as the *a priori* probability of the verb occurring in the passive also increases. The effects of this interaction are illustrated in figure 10.

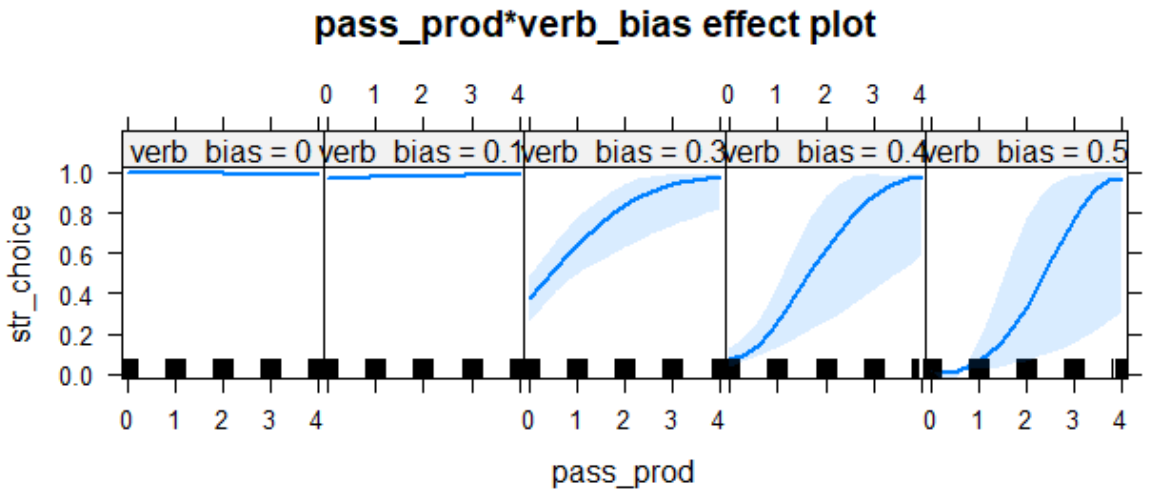

Figure 10 - Interaction between passives produced and target verb bias (passive cumulativity)

2.4. *Surprisal-sensitivity and cumulativity in the C-Oral-Brasil I corpus: discussion*

The initial predictions of higher effects of both surprisal-sensitivity and cumulativity were based largely on the staggering distribution discrepancies of the passive structure between English and Brazilian Portuguese (Guimarães and Souza, 2016; see also comparison between C-Oral-Brasil I and Penn Treebank in section 2.1 above). Following the claims of the inverse-frequency effect (Jaeger and Snider, 2007), infrequent structures tend to prime more strongly (Bock, 1986), and a structure as infrequent as the passive in BP was expected to show significantly higher effects of structural priming.

However, results from the two corpus analyses have not only failed to show differences in priming strength and decay, but they have also failed to indicate priming effects having taken place at all in the BP oral corpus. A preliminary analysis could suggest that the results from study 1 contradict the claims of the surprisal-sensitivity hypothesis and, consequently, are not accommodated by the implicit learning account of structural priming (Jaeger and Snider, 2007): the effects of prime verb passive bias on structure choice were both statistically non-significant and positive (as opposed to the tendency observed in the corpus of spoken English). Additionally, the negative correlation between prime and target verb identity would also reject the trailing activation model as a suitable account for syntactic persistence, which states that priming effects are stronger for identical lexical items (Pickering and Branigan, 2008). Finally, the positive correlation between prime and target distance and structure choice

is clearly incongruous with syntactic persistence as well as priming effects in general: since distance effects were shown to increase the likelihood of a passive target, the conclusion would be that the distance measured the number of intervening constructions between two unrelated occurrences of the passive structure.

Results from study 2 did not indicate effects of syntactic persistence either. Actives did not have an influence on structure choice, as expected for frequent structures, but neither did passives: in fact, they showed a negative correlation between cumulativity and structure choice. It is nonsensical that a negative correlation would indicate any sort of inverse effect of cumulativity, meaning that repeated activation of a structure would cause it to be less accessible. As is believed to have been the case with the positive correlation of distance in study 1, the preceding processing of passive structures appear to be unrelated to the target passives. The two predictors from study 2 whose interaction correlates positively (and are in line with observations by Jaeger and Snider, 2007) are target verb bias and number of passives produced within-speaker, which suggests that the choice of structure is based largely on existing passive distributions in Brazilian Portuguese.

While the results from the corpus analyses of Brazilian Portuguese do not add substantially to the discussion of the nature of syntactic persistence (as a result of implicit learning or trailing activation), they strongly suggest that the low productivity of the passive structure stems from employment restrictions that are very unlikely to be purely syntactic. All verbs in the corpus except for one are active-biased (cf. Appendix 1), which, according to the inverse-frequency hypothesis, would yield strong priming effects. The very fact that it did not is indicative of the existence of possibly pragmatic or even item-specific conditions that only a qualitative analysis of the structure in BP would be able to elucidate. In fact, Guimarães and Souza (2016) offer a discussion of the alternatives to the passive structure in BP, and how a view of the structure as a construction (cf. Goldberg, 1995) rather than a transformational phenomenon better accommodates its behavior in the language, and how these alternatives affect the distribution of the canonical passive.

It falls out of the scope of this study to analyze the behavior of the passive in BP, although it has become apparent that such an investigation is highly needed. These results, however, are informative enough to indicate a significant difference of the behavior of the structure between English and BP, the languages involved in the

bilingualism discussion being conducted. From the assumption of shared representations between L1 and L2 and the glaring difference of the passive construction attested in the reported corpus studies, there is sufficient information to predict a different behavior of L1 BP L2 English high-proficiency bilinguals towards the passive in the L1, which is examined by study 3.

2.5. *Study 3*

2.5.1. Design: contributions from Bock (1986) and Guimarães (2016)

The experiment in study 3 was largely based on the picture description task reported by Bock (1986, experiment 3) and Guimarães (2016). In her seminal paper, Bock (1986) manipulated priming effects of datives (double object and prepositional object) and transitives (active and passive) on a picture-description task disguised as a running recognition task. This design differed from experiments 1 and 2 (Bock, 1986) in the sense that, unlike the first 2, each prime in experiment 3 had a chance of appearing later on in the task, forcing subjects to fully process all primes in order to perform well in the memory task. In addition to the passive and active primes paired with their counterpart images, Bock (1986) manipulated the position of the agent (balanced between left and right) in experiment 3, as attempt to elucidate the absence of priming effects for human agent events in experiments 1 and 2.

In fact, Bock (1986) found an increase in the number of passives produced as a consequence of the position of both human and non-human agents. Following Bock (1986), the position of the agent was also controlled in the oral sentence elicitation experiment reported by Guimarães (2016). A group of high-proficiency L1 BP L2 English bilinguals and a group of BP monolinguals gave oral descriptions in L1 BP of images depicting transitive events from one of two lists of 24 items. Each list was comprised of 12 images showing the agent on the left and the other 12 showing the agent on the right, so that each transitive event portrayed occurred in both configurations, but were not repeated on either list – i.e. an image showed the agent on the left on one list and on the right on the other. The agents from each set of images were equally divided between human agents and non-human animal agents, with patients of equal animacy.

The absence of inanimate agents or patients in Guimarães (2016) intended to eliminate any animacy biases from event structures (animate agents tend to be

subjects more often than inanimate agents), leaving the choice of structure to each subject's wholistic interpretation of the event (Griffin and Bock, 2000) and existing representational distributions, since there was no priming or attentional manipulation of any kind (as there was in the studies by Bock (1986) and Gleitman et al. (2007), for instance). Unlike Bock (1986), Guimarães (2016) found no effects of agent position on the choice of structure in neither the high-proficiency bilingual nor the monolingual experimental groups. All the images in Guimarães (2016) presented the same color, size, and drawing style in order to control for effects of participant visual saliency that might influence subject selection and, consequently, the choice between active and passive (Griffin and Bock, 2000).

Study 3 was designed following the running recognition task from experiment 3 in Bock (1986), using the 24 images depicting transitive events from Guimarães (2016). The design and stimuli were chosen based on the findings from both studies: Bock (1986) attested that the task procedures were strong enough to yield the manipulation effects intended without exposing the nature of the study, and Guimarães (2016) was able to elicit oral descriptions that depicted the events being shown without interference of agent position or animacy.

2.5.2. Predictions

Based on accounts by Pickering and Branigan (1998), Jaeger and Snider (2007), Guimarães (2016), and Bernolet et al. (2013), we are able to make four main predictions. First, we expect that monolinguals will be primed more strongly than bilinguals. Due to the adjustment of the distributional properties of the passive structure in the bilinguals' linguistic system caused by experience with the L2, the structure is significantly more infrequent for monolinguals (Guimarães, 2016; Guimarães and Souza, 2016). According to the inverse-frequency effect hypothesis (Jaeger and Snider, 2007), the difference in passive distributions between the two linguistic profiles will result in different strength of priming effects, since infrequent structures tend to prime more strongly (Bock, 1986; Chang et al., 2006; Jaeger and Snider, 2007; Jaeger and Snider, 2013). Second, identical prime and target verbs are expected to show stronger priming effects than different prime and target verbs due to effects of lexical boost caused by residual activation of both the combinatorial node and the link between the node and the verb itself (Pickering and Branigan, 1998). Third, we predict

that the effects of syntactic persistence on low-proficiency bilinguals will be similar to the effects expected on the monolinguals, since they have not yet acquired the cognitive automatization and shift from explicit to implicit knowledge as have high-proficiency bilinguals (Bernolet et al., 2013; Hustijn, 2015). Fourth, the number of passives produced will have a positive correlation with the choice of structure, as observed in Jaeger and Snider (2007) and Bock et al. (2006). Jaeger and Snider (2007) defend that cumulativity effects are a result of speakers tracking the distribution of structures in a given context, which predicts that syntactic persistence effects increase as the number of prime sentences processed.

Therefore, we predict that four predictor variables will influence the choice of structure and verb in the descriptions:

- Linguistic profile: high- and low-proficiency L1 BP L2 English bilinguals and BP monolinguals;
- Prime type: passive and active;
- Lexical identity: the verb in the event depicted by the image matches the verb from the prime sentence;
- Previous passive structures: the number of passives structures previously produced by the subject until the occurrence of a passive structure in the description.

2.5.3. Participants

Eighteen adults between 24 and 45 years of age participated in the experiment, divided in three experimental groups: BP monolinguals, low-proficiency L1 BP L2 English bilinguals and high-proficiency L1 BP L2 English bilinguals. All participants were native speakers of BP, Brazil, and had finished high school. The Vocabulary Levels Test, or VLT (Nation, 1990), was used to classify bilinguals as low- or high-proficiency. The VLT is a test designed to measure the speaker's vocabulary knowledge through five levels matching words to descriptions. Correct matching of the 18 words of each level indicates knowledge of the 2,000, 3,000, 5,000, university level, and 10,000 most frequent words in the English language. In this study, the threshold between low- and high-proficiency was 72 points (80%) in the VLT (cf. Souza et al., 2015, Soares-Silva, 2016), while monolinguals were self-declared.

2.5.4. Material

The stimuli were comprised of 60 prime sentences in BP, each immediately followed by their 60 target images. Experimental items consisted of prime-target pairs of audio sentences and corresponding images: 12 experimental passives, 12 experimental actives, 12 filler intransitives, 6 control unlicensed double-objects, and 6 control prepositional objects (sentences 10-14, respectively). The images were all black-and-white drawings from Guimarães (2016) presented on 12cm by 12cm cards, and the sentences were recorded by a 30-year-old female BP native speaker at normal speed and articulation. Half the images of each type presented events that were more precisely described using the verb from the prime sentences, to control for lexical boost (Pickering and Branigan, 1998; Hartsuiker et al., 2008; Bernolet et al., 2013).

10. ***A mulher está sendo empurrada pelo bêbado na rua.***
    The woman is being pushed by the drunk on the street
    'The woman is being pushed by the drunk on the street.'
11. ***Um adolescente assaltou o síndico do prédio.***
    A teenager mugged the manager of building
    'A teenager mugged the building manager.'
12. ***O atleta correu durante três horas.***
    The athlete ran for three hours
    'The athlete ran for three hours.'
13. ***A mulher mostrou seu marido a casa que ela gostava.***
    The woman showed her husband the house that she liked
    'The woman showed her husband the house that she liked.'
14. ***A menina mostrou o machucado para sua mãe.***
    The girl showed the bruise to her mother
    'The girl showed the bruise to her mother.'

Figure 11 - Image used in the event "push"

Figure 12 - Image used in the event "mug"

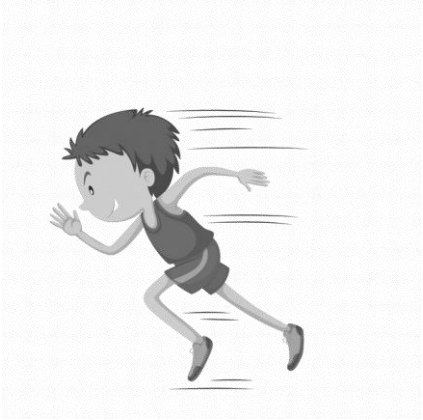

Figure 13 - Image used in the event "run"

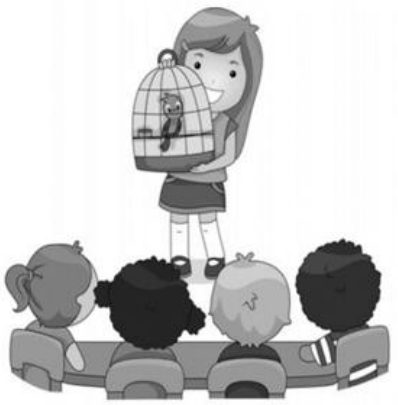

Figure 14 - Image used in the event "show"

Each trial was pseudorandomized so that all primes immediately preceded their targets and prime-target pairs occurred in a different order every trial. No two prime-target pairs of the same type were presented in sequence, and the ditransitive sentences and the intransitive images were doubled to serve as the targets for the cover running recognition task (cf. Bock, 1986), with a total of 24 repetitions in a set of 120 stimuli.

2.5.5. Procedures

The instructions for the primary task of recognition were for subjects to pay close attention to sentences and images, as they were supposed to indicate whether each item was new to the set or had already been presented. The secondary tasks of sentence repetition and picture description took place before subjects gave their answers to the primary task. Thus, upon presentation of a sentence, subjects repeated it out loud, verbatim, and only then emitted their judgment for the recognition task. Likewise, subjects described the images in as much detail as possible and then indicated whether they recognized the image. After the end of the whole task, subjects were asked if they could identify the purpose of the study to control for any learning effects from the pseudorandomization.

2.5.6. Voice alternation data

A total of 1080 descriptions were collected across the three experimental groups. The 216 repetitions were eliminated from the study, because they did not constitute controlled prime-target pairs. The 72 trials related to the ditransitive verbs *trazer* (bring) and *vender* (sell) were also eliminated because subjects failed to identify

the events portrayed in the target images in more than half of the descriptions, indicating that any effects on those trials could not be associated with verb frequencies or prime-target identity. A total of 90 individual trials where subjects also failed to identify the event were eliminated for the same reason, but without consequences to the sentence category to which they belonged.

The voice alternation trials resulted in 372 descriptions. Structures chosen included actives, passives, intransitives, compound NP subjects, noun phrases, oblique objects (not passivizable in BP), and prepositional phrases, shown in table 7:

Structure	*Occurrences*	*Example*	*Verb*
active	312	Homem abanando a mulher.	abanar (fan)
passive	30	A Cleópatra sendo abanada por um súdito.	abanar (fan)
oblique object	13	O menino olhando para a menina de binóculo.	espiar (peek)
compound NP subj.	8	Duas mulheres se beijando no rosto.	beijar (kiss)
NP	5	Um assalto.	assaltar (rob)
intransitive	2	A moça sentada com índio abanando.	abanar (fan)
prepositional dative	2	O homem abanando a folha de alguma planta numa mulher.	abanar (fan)

Table 7 - Structures used in voice alternation descriptions

The responses were categorized based on the following conditions:

- Actives: presented a direct NP complement

  15. *Homem abanando* ***a mulher.***
      Man fanning **the woman.**

- Passives: presented the verb *ser* (be) followed by the participle of the main verb, with or without the agentive by-phrase

  16. *A Cleópatra* ***sendo abanada por um súdito.***
      The Cleopatra **being fanned by a subject.**

- Oblique objects: presented an NP complement preceded by a preposition

  17. *O menino olhando* ***para a menina de binóculo.***
      The boy looking **at the girl with binoculars.**

- Compound NP subjects: presented both participants in the subject and no verb complement

  18. ***Duas mulheres*** *se beijando no rosto.*
      Two women (refl.) kissing on the cheek.

- Noun phrases: did not present a finite verb

  19. *Um* ***assalto.***
      A **robbery.**

- Intransitives: did not present a verb complement

  20. *A moça sentada com índio* ***abanando.***
  The woman seated with Indian **fanning.**

- Prepositional datives: presented both NP and PP complements

  21. *O homem abanando* ***a folha de alguma planta numa mulher.***
  The man fanning **the leaf of some plant on a woman.**

It is important to emphasize that the classification of verbs that require complements as intransitive follows the claim of syntax of oral production that there needs to be an acoustic signal for a complement to be considered (Raso and Mello, 2012). In sentence (20), even though it is inferable from the connection to the image (figure 15) that the “Indian" mentioned is in fact fanning the woman, the linguistic expression in itself does not allow us to affirm that the there is a complement for that verb. One can note that, if the image is taken away, the question “*who is the Indian fanning?*” becomes entirely plausible. For the purposes of this analysis, all classifications except for actives and passives were combined into a third category henceforth named “other”, amounting to 30 occurrences.

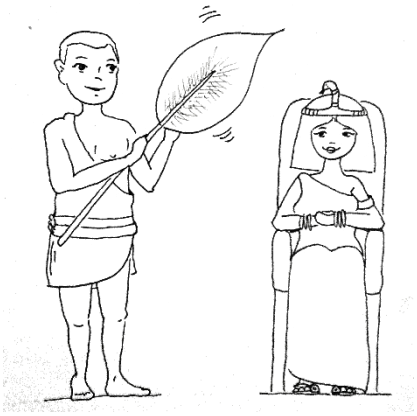

Figure 15 - Image used in the event "fan"

2.5.7. Results

The data was analyzed using logistic regression for the categorical response variables target type (passives or non-passives) and choice of verb – note that choice of verb refers to the verb from the *description* matching the verb from the prime sentence, while lexical identity refers to the *image* depicting the same event as the prime sentence. First, the entire data set was analyzed to investigate whether lexical identity had an overall effect on choice of verb, that is, if subjects chose the same verb

as the one from the prime sentence when the image depicted the same event. There was no significant effect of lexical identity on subjects' choice of verb (Z = -0.034, *p* = 0.973). The same analysis was performed for the voice alternation data set (only active and passive primes and images), with equally non-expressive results (Z = -0.014, *p* = 0.989; figure 16). Interestingly, lexical identity had a significant negative effect on target type for passive primes (Z = - 2.456, *p* = 0.014; figure 17).

Given the inexpressive influence of lexical identity on subjects' choice of verbs or target types, the issue then becomes whether the actual production of a description using the verb from the prime sentence favors the production of passive structures. In this analysis, the response variable choice of verb (identical or different from the verb in the prime sentence) becomes the predictor variable for target type, with the identical

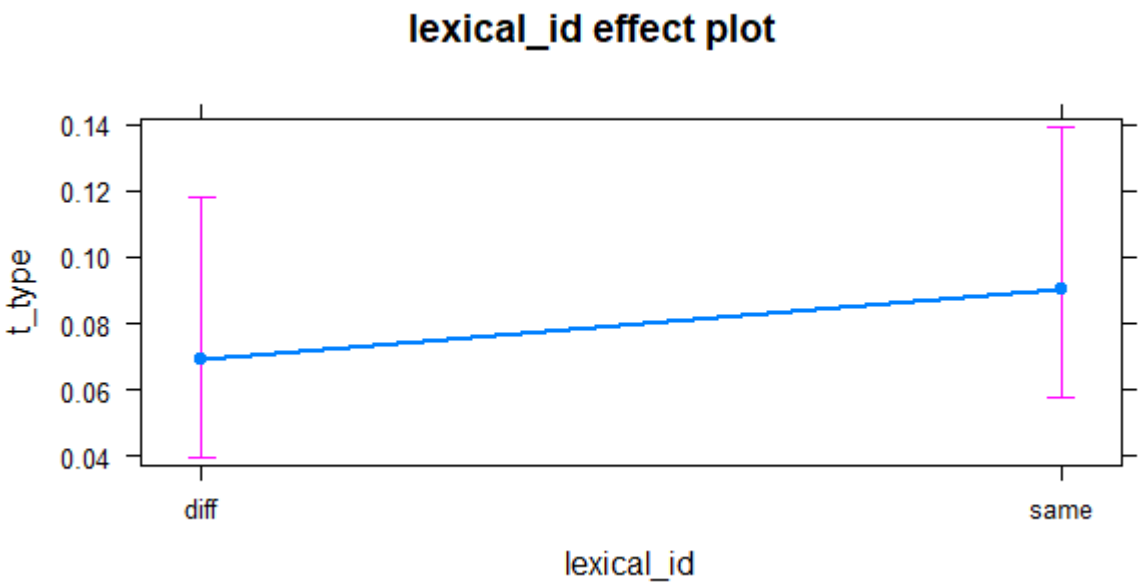

Figure 16 - Effects of lexical identity on structure choice

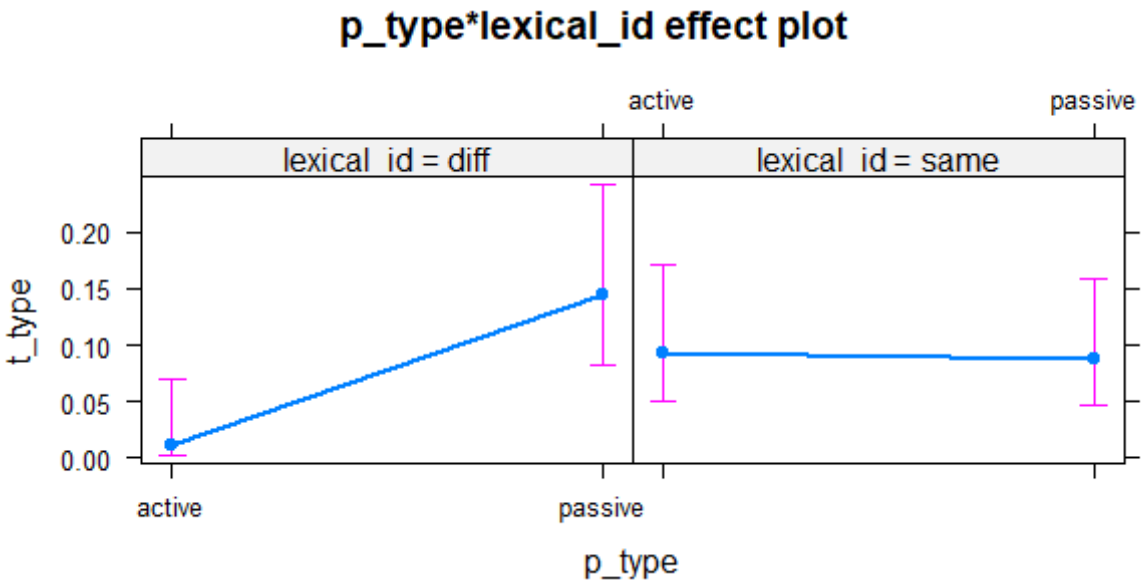

Figure 17 - Interaction between prime type and lexical identity

condition as the level of reference. The choice of an identical verb did not have significant effects on target type (Z = 0.420, *p* = 0.67436). Interactions between choice of verb and either passive primes (Z = 0.015, *p* = 0.9881) or linguistic profile (monolinguals, low- and high-proficiency bilinguals) failed to show significant effects (table 8):

	Estimate	*S.E.*	*Z value*	*p value*
(Intercept)	-1.386	1.118	-1.240	0.215
verb ID * low-bi	17.065	1318.728	0.013	0.990
verb ID * monolingual	18.221	1398.722	0.013	0.990

Table 8 - Interaction between choice of verb and profile

Production of passive structures in the descriptions in this study was in sharp contrast to what was observed in Guimarães (2016). As opposed to the extremely few passives produced by monolinguals in Guimarães (2016) – 3.75% of monolinguals' descriptions were passives, as opposed to bilinguals' 11.41% of passive structures – monolinguals in this experiment produced a significantly higher number of passives than bilinguals in general (Z = 2.225, *p* = 0.0261). This difference can be attributed to the fundamental difference in the task: Guimarães (2016) manipulated agent position

(left or right, with no significative results), while the present design manipulated priming of passive structures.

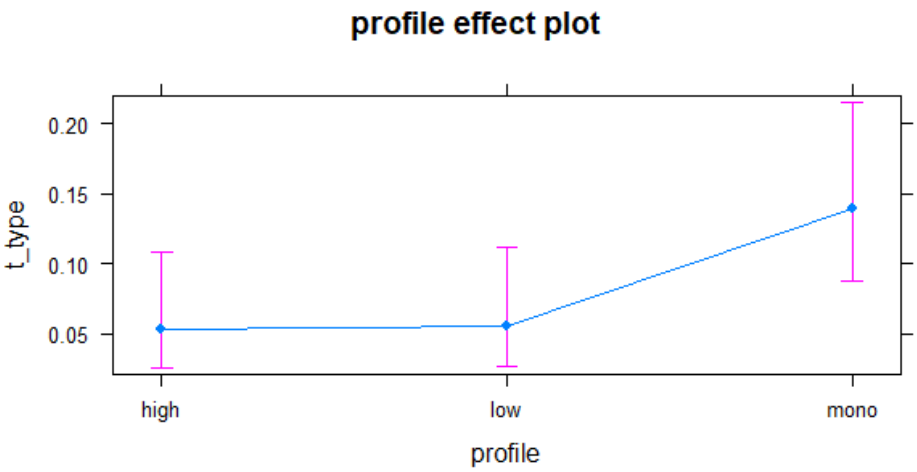

Figure 18 - Production of passives by linguistic profile

In fact, there were effects of prime type on target type (Z = 2.073, *p* = 0.0382), indicating that passives tended to occur more after other passives than after actives. This is suggestive that, unlike the conversations from C-Oral-Brasil I, the image descriptions were influenced by the structure of the prime sentence. The interaction of interest, however, is between prime type and target type within each of the linguistic profiles. Interestingly, although monolinguals did produce more passives than bilinguals, the effects of passive primes were not significant across the three experimental groups (table 9).

Interaction	*Estimate*	*S.E.*	*Z value*	*p value*
(Intercept)	-3.0910	0.5902	-5.237	1.63e-07
pass prime * low-bi	-0.6205	1.1102	-0.559	0.576
pass prime * monolingual	1.2726	1.0328	1.232	0.218

Table 9 - Interaction between prime type and profile on choice of structure

Finally, the number of passives previously produced was positively correlated with the target type across the voice alternation set (Z = 3.301, *p* < 0.001, figure 19). Similarly to the prime type effect, cumulativity did not vary across the three linguistic profiles.

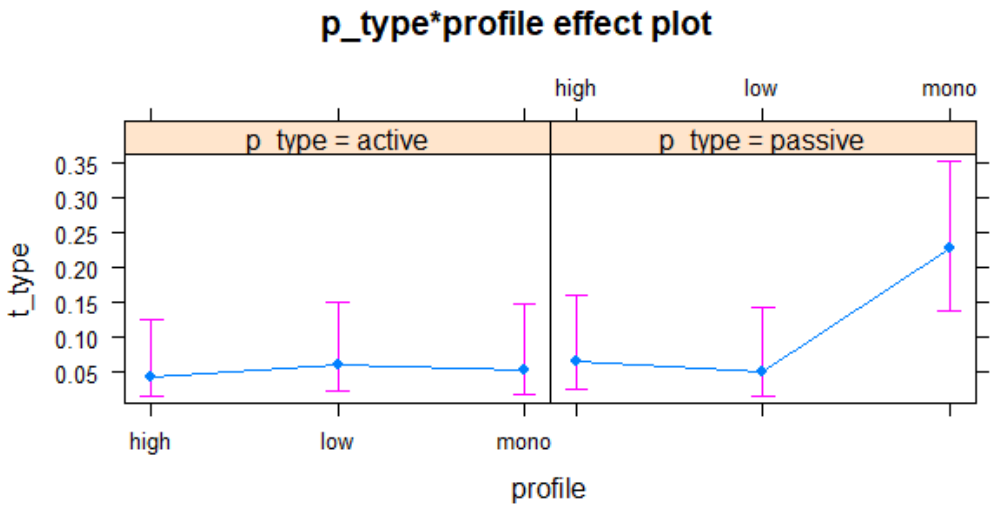

Figure 19 – Interaction between prime type and linguistic profile

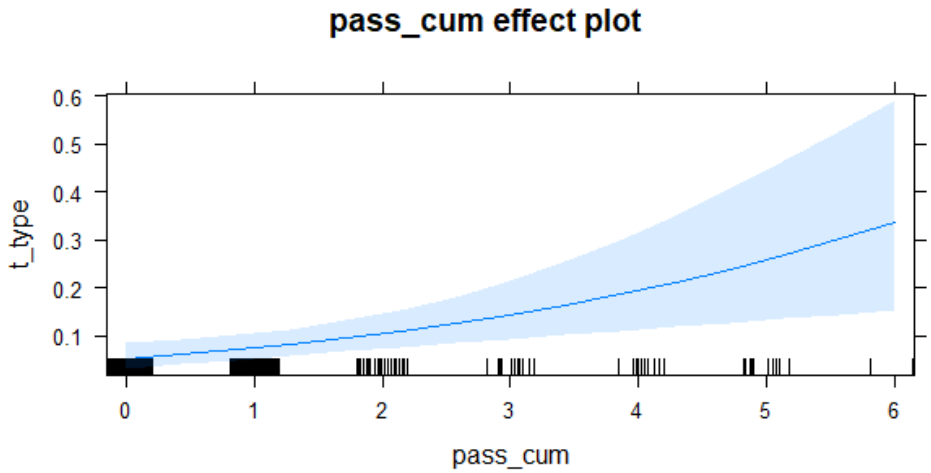

Figure 20 - Passive cumulativity on choice of structure

2.5.8. Discussion

Four predictions were made in the design of this study: that monolinguals would be primed more strongly than bilinguals; that lexical identity would increase priming effects; that low-proficiency bilinguals would behave similarly to monolinguals; and that previously produced passives would have positive effects on the choice of subsequent structures.

The first prediction was partially confirmed: although monolinguals produced more passives than bilinguals, the priming effects observed among subjects from this group did not differ significantly from what was observed in the bilingual groups. This

indicates that subjects of all linguistic groups produced more passives after hearing prime sentences in the passive than after hearing active primes, confirming the occurrence of structural priming in the task. As the images in study 3 were extracted from the experiment in Guimarães (2016) and subject classification followed the same parameters (bilinguals were considered highly proficient if they scored more than 80% in the VLT), it is possible to use the results from the oral sentence elicitation task as a baseline to analyze the present priming effects. Table 10 brings a comparison between descriptions in Guimarães (2016), under the column “free production”, and the results from study 3, under the column “primed production”:

Profile	*Free production*		*Primed production*	
	passives	*non-passives*	*passives*	*non-passives*
bilinguals	21	163	7	124
monolinguals	6	155	16	99

Table 10 - Production in free and primed tasks

A *post-hoc* chi square analysis shows that the rate of passive production did not vary significantly for bilinguals as a function of the task type ($\chi^2(1) = 3.4807$, $p = 0.0621$). On the other hand, the number of passives produced by monolinguals increased significantly: $\chi^2(1) = 9.4888$, $p = 0.0021$. This discrepancy supports the interpretation that speakers tended to describe images using the passive structure after hearing a sentence in the passive.

Although the difference in the strength of priming between monolinguals and the bilingual groups was not statistically significant, it signals a slight effect of magnification of priming effects for the monolingual group that explains the significant variation in this group’s performance between free and primed production. However, the group of high-proficiency bilinguals showed a tendency to avoid the passive in the primed task in comparison to the free task: passives were produced 7% less in this study. Given that prime type influenced target type in the voice alternation data set, it would be expected that high-proficiency bilinguals also showed increase the number of passive structures produced, not a decrease. This raises the issue of what may have caused their contradictory behavior. The data only allows for speculation as to what constrained the use of passive by bilinguals at this point, but a more robust pool of subjects could potentially boost the difference in both strength of priming between

experimental groups and the difference observed in bilingual production. Nevertheless, both corpus and laboratory findings indicate that there may be (quite possibly pragmatic) constraints on the passive in BP in comparison to English that cause bilinguals to tend to reject the structure upon hearing it, and monolinguals to tend to reject it in unprimed conditions.

Overall, passives structures primed monolinguals to produce subsequent passives more often than they do so under normal circumstances. The fact that monolinguals' production suffered positive interference from priming while bilinguals' production decreased compared to the baseline is in itself support for the inverse-frequency effect of surprisal. In the case of BP, where passives are extremely infrequent, monolinguals shifted from approximately 4% to 16% rate of production of passives. The comparison between linguistic profiles was not so expressive. Figure 19 clearly shows a higher tendency by monolinguals to produce passives after the prime in relation to the bilingual groups, but the lack of statistical robustness prevents us from stating categorically that knowledge of L2 English is the cause of differences in surprisal-sensitivity of passives in BP.

The second prediction was not confirmed: there were no overall effects of lexical boost (as stated by Pickering and Branigan, 1998), despite the odd negative correlation between lexical identity and target type when the primes were in the passive. Neither did lexical identity (identical events in prime sentence and target image) predict the choice of verb in the description, nor did identity between verbs in the passive prime sentence and the description predict choice of passives in the target. The distinction between the lexical identity and choice of verb comes from the observation that many lexically identical images were described using synonyms of the verb in the target sentence, while expressing the event intended by the image. The image portraying the event *perseguir* (chase, figure 21), for instance, was described with the verb phrase *correr atrás* (run after) in 67% of the descriptions.

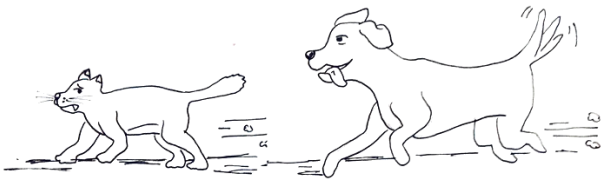

Figure 21 - Image used in the event "chase"

Implicit learning accounts of structural priming assume that lexical boost effects are not an indication of activation of implicit memory; instead, they are said to be caused by explicit memory of the surface structure of the target. In fact, explicit memory might be the construct behind the significant negative effect observed in the interaction between lexical identity and target type. As the only experimental group that decreased the number of identical verbs between prime sentence and description was that of the high-proficiency bilinguals, it is possible to hypothesize that, upon encounter with an image directly related to the passive prime sentence, the subjects from this group avoided verbatim repetition and intentionally changed the verb in the target description. Once again, the motivation behind this apparent rejection of passives by high-proficiency bilinguals under priming conditions needs further investigation.

The third prediction of this study concerned the performance of the low-proficiency bilingual group. From the assumption that structural representations and their distributional properties are not yet shared between the L1 and the L2 in this level of proficiency, priming effects were expected to be of similar strength within the low-proficiency and monolingual groups. Nonetheless, the only suggestive (but not significative) distinction observed was between monolinguals and bilinguals in general, whose overall performance was similar. Bernolet et al. (2013) argue that low-proficiency mitigates effects of cross-linguistic structural priming while lexical boost effects are stronger; possibly due to the low-proficiency speaker's dependency on item-specific representations. Conversely, study 3 failed to observe either lexical boost or L2 proficiency effects on structural priming.

In this study, the activation of the passive structure representation came from the L1 instead of the L2, which is the main point of distinction study 3 and Bernolet et

al. (2013) in the sense that proficiency was herein taken as a predictor of performance because of distributional properties rather than access to the lemma level. Regardless of the difference between theoretical standpoints concerning the nature of structural priming (implicit learning or residual activation) taken in this study and theirs, the failure to observe L2 proficiency interference on the expected priming effects does not entail that representational sharing takes place in early stages of L2 proficiency. Instead, we argue in favor of an early sharing of the passive structure as a result of facilitation due to the structure's superficial similarity in L1 and L2. It appears that morphosyntactically identical structures from the L2 are abstracted into the procedural memory earlier than similar (e.g. the genitive in Dutch and English, cf. Bernolet et al., 2013) or unlicensed ones (e.g. the induced movement alternation in English, cf. Souza et al., 2014), possibly because the speaker is able to generalize and abstract rules from the structure's distributional properties after fewer instances of L2 experience.

Finally, the prediction that an increase in the number of passive structures produced previously would increase the likelihood of a passive being produced in the target was confirmed. Differently from surprisal-sensitivity, cumulativity showed clear effects on the production of target passives, in line with implicit learning accounts that tracking structure distributions increases the likelihood of choosing it over its alternatives given the context.

3. GENERAL DISCUSSION

3.1. *Structural priming as learning*

The architecture of implicit learning processes outlined in the dual-path model is based on adjusting predictions to error (Chang et al., 2006). The speaker hears a word and predicts the next based on prior knowledge of the state of the language. If the subsequent word does not correspond to the one predicted, the entire system adjusts its distributional weighs via backpropagation: a new, arbitrary association is made between the word previously heard and the following unpredicted word. The prediction cycle is then resumed, using the last heard word as the input for prediction of the next. This serial word processing mechanism is able to abstract lexical and syntactical categories from sentences because of its connection to the message associated with the string of words comprehended. The message system attributes event roles to concepts on the basis of activation, with more prominent roles being assigned to more prominent concepts. It is the role assignment that determines the structure of the sentence.

Griffin and Bock (2000) defend that events are entirely comprehended before the onset of speech. Thus, the selection of a subject of a transitive event, for instance, is not determining of the apprehension of the event as a whole, but a consequence of activation prominence. Findings from Gleitman et al. (2007) support the notion of structure selection trough activation levels (Chang et al., 2006) and the perspective of wholistic apprehension prior to speech production in Griffin and Bock (2000). Manipulation of speakers' attention to one or the other participant of the event forced the production of less frequent structures; nevertheless, the eye movements of the speakers prior to the fixation on the sentence subject indicated that the entire event had been apprehended before the descriptions were made.

It is important to stress that concept activation is not a fundamentally automatic mechanism, but is in fact amenable to visual and auditory stimuli (Levelt et al., 1999; Segalowitz and Hulstijn, 2005). Therefore, the decision between describing the event in figure 22 as sentence 22 or 23 is based on whether the activation levels for the concepts LIGHTNING or HOUSE are higher:

22. The lightning is striking the house.
23. The house is being struck by lightning.

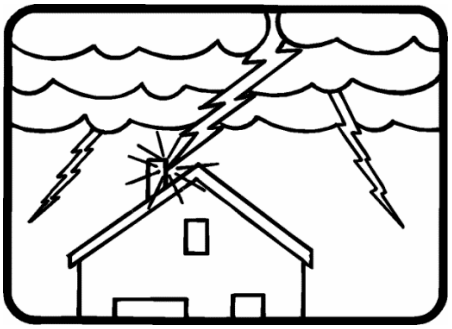

Figure 22 - Image portraying the event "strike"

Knowledge of the passive structure (or any structure) comes from the syntactic abstractions acquired by the prediction network. Thus, as the speaker attributes the role of patient to the most prominent concept, the production system activates the appropriate representations that result in a grammatical sentence.

Given that speakers are constantly adjusting their prediction expectations during language processing, it follows that every episode of language use results in learning in some level. In this account, structural priming is a result of learning in that the tendency to generalize recently processed structures to different utterances is caused by the adjustment of the distributional weights of the structure upon its processing. Priming effects derived from distributional reconfigurations have been found to be long-lived and, therefore, cannot rely on recent activation alone (Chang et al., 2000). The longevity and strength of priming effects as a function of structure frequency is referred to as the surprisal-sensitivity hypothesis, where surprisal is defined as the log inverse of the item's frequency. Structural priming has been also found to be cumulative, as a larger number of preceding structures provide speakers with a more accurate estimate of its probability given a context (cf. Jaeger and Snider, 2007, and the results from study 3).

3.2. *Distributional learning in late bilingualism*

Inferring abstract rules from distributions is a process that is not exclusive to infants in first language acquisition, but a domain-general mechanism that is intrinsic to cognitive processes including late bilingualism. We argue that processing mechanisms of L1 BP L2 English high-proficiency bilinguals are porous to linguistic episodes in both L1 and L2, based on the constructs of L2 proficiency, implicit learning, and distributional learning mechanisms.

Aslin and Newport (2012) define statistical learning as “a mechanism that enables adults and infants to extract patterns of stimulation embedded in both language and visual domains” (p. 170). Saffran et al. (1996) observed 8-month-old infants’ abilities to detect word boundaries from transitional probabilities of syllables (transitional probabilities within words are higher than between words), even in the absence of other informative cues such as pauses or intonation. Their findings supported the sufficiency of statistical cues in abstracting rules from distributions in the input. Further studies by Aslin et al. (1998) and Aslin and Newport (2012; 2014) attested that this mechanism is modality-, domain-, and species-general, taking place in different linguistic processes, in different paradigms such as language, images, and music, and among human as well as non-human species.

The construct of L2 proficiency is directly connected to processes of automatization and shift from explicit to implicit memory (Ullman, 2004; Schneider and Shiffrin, 1977; Bernolet et al., 2013), which are sensitive to frequency effects from experience with the L2. Thus, a high-proficiency bilingual is believed to share abstract structural representations between the L1 and L2 to the extent that their distributional properties merge between the languages (Hartsuiker et al., 2004; Bernolet et al., 2013, Guimarães, 2016, Souza and Oliveira, 2014). A high-level of proficiency is the condition for the sharing of linguistic prediction system (cf. Chang et al., 2006) as it presupposes that knowledge and domain of the L2 are automatic and implicit enough for structures to be stored as abstract rules, rather than explicit item- and language-specific representations.

The studies conducted in this dissertation were designed to investigate the assumptions that underlie the general hypothesis. First, it was necessary to examine structural priming effects in BP in naturalistic conditions to determine whether the distribution discrepancy of the structure in BP and English observed by Guimarães (2016) indeed affected the properties of surprisal-sensitivity and cumulativity, widely reported in the literature (Bock, 1986; Chang et al., 2000; Chang et al., 2006; Jaeger and Snider, 2007; Jaeger and Snider, 2013, among others). The results from the analysis of the corpus of spoken BP C-Oral-Brasil I (Raso and Mello, 2012) were inconclusive with respect to these properties of structural priming, since the effect did not seem to have taken place in the dialogues and conversations analyzed. Particularly, the backwards correlation observed between prime and target distance

and choice of structure was a clear indication that, rather than measuring the decay of syntactic persistence, these numbers (which ranged from 1 to 114) were simply indications of the number of intervening actives between two unrelated instances of passive structures. The cumulativity analysis was in accordance with the interpretation that priming did not occur, as the only significant factor was the interaction between target verb bias and passives produced – two measures that reflect within-speaker effects on language production.

The second assumption under examination was that proficiency constrains distributional learning over structures from the L2, as L2 proficiency implies a high level of automaticity in L2 processing and reflects a state of shared structural representations between L1 and L2. As L2 proficiency increases, structures learned from the L2 depart from explicit memory as item- and language-specific to abstract representations whose distributional properties can be generalizable over the linguistic system as a whole. As an attempt to account for the potential proficiency threshold for representational sharing, study 3 divided its subjects into groups of monolinguals, low-proficiency and high-proficiency bilinguals, so that the performance of low-proficiency bilinguals was compared to those of monolinguals and high-proficiency bilinguals in terms of surprisal-sensitivity. The inverse-frequency effect predicted a sharp difference in priming magnitude between the monolingual and high-proficiency groups due to L2 influence on bilinguals' passive distributions, but not between monolingual and low-proficiency groups, under the assumption that representations are not yet shared between L1 and L2.

Results from study 3 contradicted both predictions: neither did priming effects for monolinguals differ significantly, nor did low-proficiency bilinguals' performance resemble that of monolinguals. Although the rate of passive production among monolinguals was significantly higher in comparison to bilingual groups in the priming task as well as to monolinguals in the free production task (Guimarães, 2016), it was not possible to state that linguistic profile was determining of passive structural priming. Nevertheless, the data allows for speculation concerning the role of L2 proficiency on the mechanisms of language production.

3.3. *Similarity modulation on shared representations between L1 and L2*

The passive is mophosyntactically identical in BP and in English: it consists of a copula verb followed by the participle of the main verb and an optional agentive by-phrase. Consequently, the difference in its distributional properties (cf. Guimarães and Souza, 2016) must be attributed to language-specific constraints other than syntax. The increased production of passives by bilinguals in relation to monolinguals in Guimarães (2016) strongly suggests that these constraints are abstracted from the structure's distributional properties and, therefore, shared between L1 and L2. Results from study 3 indicate that generalization over passive structure distribution takes place in early L2 proficiency stages.

It appears that morphosyntactic similarity facilitates distributional learning in the L2 for both syntactic and semantic-pragmatic properties of structures. Whereas strength of structural priming did not differ from low- to high-proficiency bilinguals in this study, Bernolet et al. (2013) found that L2 proficiency directly constrained priming magnitude and lexical boost. Their study was based on the structural representation of the genitive, a similar – but not identical – structure in English and Dutch. The latter language limits the Saxon genitive to proper names and nouns of specific reference (e.g. father), and presents a pronominal alternative in spoken language that is morphosyntactically different from its English counterpart[7]. Their results provided evidence in support of a timeline of structure abstraction throughout the development of L2 proficiency, reflected on its direct correlation to strength of structural priming and negative correlation to lexical boost effects.

Similarity facilitation becomes more apparent in cases of structures from the L2 that are unlicensed in the L1. Souza et al. (2014) report that high-proficiency bilinguals accept L1 BP sentences presenting unlicensed structures (namely, the caused motion alternation and the resultative structure) from L2 English significantly more than BP monolinguals. The direct translation of 24 is unlicensed in BP (sentence 25), and the proposition expressed by the verb *run* in 24 could only be acceptably expressed in BP as the periphrastic causative in 26.

24. The coach ran the students around the field.
25. * O treinador correu os alunos em volta do campo.

[7] For a detailed comparison of genitives in English and Dutch, see Bernolet et al. (2013), p. 290-291.

26. O treinador fez os alunos correrem em volta do campo.
'The coach made the students run around the field.'

While the caused motion alternation does not result in a licensed structure when directly translated into BP, the translation of the resultative construction results in a licensed structure whose meaning differs from the original. In sentence 27, the adjective *clean* is the result of the action *wipe*, whereas in sentence 28 the corresponding adjective *limpa* is a modifier of the object *mesa*. Thus, sentence 28 does not convey the meaning intended in 27.

27. Samuel wiped the table clean.
28. Samuel esfregou a mesa limpa.

The higher acceptability of unlicensed structures such as in 25 and 28 is indeed compelling evidence of the learnability of structures from the L2 whose L1 counterparts are unlicensed. Representational sharing of these structures, however, appears to be modulated by morphosyntactic similarity. While the structural BP counterparts of English resultatives are acceptable and fairly simple active constructions (e.g. sentences 28), caused motion alternation sentences benefit from a similar synthetic causative possibility in a BP variant productive in the state of Minas Gerais, home state of the majority of the subjects:

29. ***A professora correu o menino para fora da sala.***
The teacher ran the boy to out of the classroom.
'The teacher ran the boy out of the classroom.'

Ciríaco (2007) states that the synthetic causative alternation in the BP variant of Minas Gerais appears to be constrained by item-specific lexical-semantic properties that license sentence 29, but not sentence 25. Similarly, the verbs *estudar* (study) and *almoçar* appear in the causative alternation when conveying a meaning similar to the construction "provide someone with":

30. ***O pai estudou os filhos até a faculdade.***
The father studied the children until the university.
'The father put his children through school and university.'

31. ***Eu já almocei os meninos.***
I already lunched the boys.
'I have already given the boys lunch.'

Morphosyntactic identity appears to facilitate processing – and acceptance, in consequence – in spite of the meaning distinction between the syntactic structure in English and in BP. Further evidence that the acceptability in comprehension owes more to morphosyntactic similarity than learnability comes from Trujillo (2018), who replicated the study reported in Souza et al. (2014) about the acceptability of the caused motion alternation by L1 Spanish L2 English high-proficiency bilinguals. Unlike BP, Spanish does not license the synthetic causative structure (sentence 32):

32. **** El capitán marchó a los soldados hasta el campamento.***
The captain marched obl. the soldier to the camp.
'The captain marched the soldiers to the camp.'

Trujillo (2018) failed to replicate the results from the caused motion experiment by Souza et al. (2014): both high- and low-proficiency L1 Spanish L2 English bilinguals considered causatives such as 32 unacceptable. Hence, discrepant levels of acceptability of the caused motion alternation from L2 English in L1 BP and L1 Spanish are indicative of speakers' sensitivity to similarity of foreign structures based on existing structures from the L1.

Let us restate the second assumption underlying the main hypothesis of this research: generalization over structures from the L2 is constrained by L2 proficiency. While there is evidence such as the results by Bernolet et al. (2013) that offer support to this assumption, the performance of low-proficiency bilinguals in study 3 suggested that, in addition to L2 automaticity, representational sharing is also modulated by how similar the new structure is to existing L1 representations. The distributional properties of the structures under examination in the within-language priming manipulation in study 3, the cross-linguistic priming manipulation in Bernolet et al. (2013), and the acceptability judgments in Souza et al. (2014) and Trujillo (2018) are believed to be the main factor in differences between and, in the case of the caused motion alternation from L2 English, within L2 proficiency groups.

On the one hand, the passive structure is morphosyntactically identical in BP and English, and its priming strength was virtually the same for both high- and low-proficiency L1 BP L2 English bilinguals (with the only exception being the apparent rejection displayed by high-proficiency bilinguals). Overall, frequency distributions from the L2 showed a tendency to magnify priming strength on BP monolinguals due to surprisal-sensitivity, not on low-proficiency bilinguals. The genitive, on the other hand, is similar in English and Dutch as both have the alternatives of the Saxon genitive (*'s*) and the pronominal possessive; however, Dutch presents stricter restrictions on the use of the Saxon genitive as well as different morphosyntax regulating the use of the pronominal genitive in spoken language. Cross-linguistic priming was stronger for high-proficiency L1 Dutch L2 English bilinguals than for the low-proficiency group, which Bernolet et al. (2013) attributed to the availability of the combinatorial node for both languages. In turn, the caused motion alternation is constrained by item-specific and diatopic properties in BP, and not at all licensed in Spanish. High-proficiency L1 BP and L2 English and L1 Spanish L2 English bilinguals differed significantly in their acceptance levels of the caused motion alternation in their L1: L1 BP bilinguals were significantly more tolerant to the unlicensed structure from L2 English than BP monolinguals, whereas L1 Spanish bilinguals and monolinguals did not accept these Spanish counterparts of the structure from L2 English.

Data from the status of these three structures present in L2 English and their distributional properties (if any) in L1 BP, Dutch, and Spanish lead us to conjecture that there is a gradience of similarity facilitation on the abstraction of structures encountered in the L2. Representations for morphosyntactically identical structures are shared in earlier stages of L2 proficiency regardless of their differences in usage distributions or language-specific pragmatic constraints, as is the case of the passive in BP and English (Guimarães and Souza, 2016). Morphosyntactically similar structures such as the genitive in Dutch and English are shared in the bilingual linguistic system later on in the development of L2 proficiency. The distinctive features of the structure are abstracted into grammatical rules as more instances of usage are processed, involving a wider range of lexical items, until the structure is generalized across languages as are identical structures (Bernolet et al., 2013). Learned L2 structures that do not have a morphosyntactic counterpart in L1 are shared between

languages only at late stages of L2 proficiency. Sentence pairs 33-35 illustrate the passive, the genitive, and the caused motion alternation and their English counterparts.

33. ***A casa está sendo atingida por um raio.***
    The house is being hit by a lightning
    'The house is being hit by lightning.'
34. ***Het meisje haar appel***
    The girl her apple
    'The girl's apple'
35. **** El capitán marchó a los soldados hasta el campamento.***
    The captain marched obl. the soldier to the camp.
    'The captain marched the soldiers to the camp.'

The facilitation provided by L1 similarity can be explained in terms of restrictions preventing computational explosion (Aslin and Newport, 2014), i.e. the overwhelming number of statistical computations that can be done from a complex set of input. The learning system focuses the statistical computations on relevant aspects of the input in order to reach specific generalizations. For instance, a bilingual must isolate verb category distributions from, say, syllable transitional properties (computed in speech segmentation). Upon encounter with a structure from the L2 that possesses a morphosyntactically identical counterpart in the L1, distributional computations on word form and argument structure, for instance, are unnecessary, and the learning mechanism can focus on distributional properties sooner. This conjecture does not necessarily assume a hierarchy of distributional computations in second language learning (although some SLA models do), but it follows the nature of the computations that result in explicit and implicit knowledge. Linguistic knowledge is first stored as arbitrary and language-specific representations relying largely on particular superficial forms rather than implicit generalizable rules.

3.4. *Late L2 learning and processing as byproducts of surprisal*

Surprisal effects in late bilinguals support the claim that second language emerges from error-driven learning as much as does first language. Every instance of exposure to structures from the L2 causes the linguistic system to adjust its predictions to accommodate the new data from that episode of language processing, regardless of the existence of morphosyntactically identical structures in the L1. Similar structures

are generalized over the entire linguistic system sooner in the course of L2 proficiency development than novel structures because speakers are able to compute their distributions from a smaller set of input, given that other superficial generalizations can be retrieved from procedural knowledge of the L1. Proficiency, then, features as a sort of time stamp of representational sharing of structures between L1 and L2, that is, the shift of novel structures into procedural memory.

The emergence of second language relies on surprisal effects inasmuch as infrequent or novel structures cause greater prediction adjustments in the linguistic system. Caused motion alternation structures, for example, were once salient to the low-proficiency bilingual regarding their meaning associations and semantic-pragmatic distributions that differ from the identical but highly restrained morphosyntactic counterpart in BP. Episodes of L2 processing modulate the processing system as a whole, since its current state at the time of encounter determines whether surprisal rates are higher, as they are for passive structure for BP monolinguals, or lower, in the case of bilinguals that have already generalized over the structure distributions as a whole from experience with the L2.

The understanding of late bilingualism as a byproduct of surprisal computations into the linguistic system finds support in both within- and between-language priming effects, which are sensitive to the state of both the L1 and the L2 since early stages of proficiency, as demonstrated in study 3. This is a promising line of research that still requires naturalistic, online behavioral and neurophysiological examination concerning the specific structure being learned, the stage of L2 proficiency of the speaker, and the current state of the two or more languages involved in order to achieve a full understanding of the learning and processing mechanisms that underlie bilingualism. So far, we can affirm with a fair level of confidence that the linguistic knowledge construction is continuously modulated by instances of use.

4. REFERENCES

ASLIN, R. N.; NEWPORT, E. L. Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science*, 21(3), 170-176, 2012.

ASLIN, R. N.; NEWPORT, E. L. Distributional Language Learning: Mechanisms and Models of Category Formation. *Language Learning*, 64(2), 86-105, 2014.

ASLIN, R. N.; SAFFRAN, J. R.; NEWPORT, E. L. Computation of Conditional Probability Statistics by 8-Month-Old Infants. *Psychological Science*, 9(4), 321–324, 1998.

BATES, D.; MAECHLER, M.; BOLKER, B.; WALKER, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48, 2015.

BELAVINA KUERTEN, A.; MOTA, M.; SEGAERT, K.; HAGOORT, P. Syntactic priming effects in dyslexic children: a study in Brazilian Portuguese. Poster presented at the 22nd Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP 2016), Bilbao, Spain, 2016.

BERNOLET, S.; COLLINA, S.; HARTSUIKER, R. J. The persistence of structural priming revisited. *Journal of Memory and Language*, 91, 99-116, 2016.

BERNOLET, S.; HARTSUIKER, R. J.; PICKERING, M. From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, 127, 287-306, 2013.

BOCK, J. K. Syntactic persistence in language production. *Cognitive Psychology*, 18, 355-387, 1986.

BOCK, J. K.; GRIFFIN, Z. M. The Persistence of Structural Priming: Transient Activation or Implicit Learning? *Journal of Experimental Psychology: General*, 129(2), 177-192, 2000.

BURNS, T. C.; YOSHIDA, K. A.; HILL, K.; WERKER, J. F. The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, 28, 455-474, 2007.

CHANG, F.; DELL, G. S.; BOCK, J. K. Becoming Syntactic. *Psychological Review*, 113(2), 234-272, 2006.

CHANG, F.; DELL, G. S.; BOCK, J. K.; GRIFFIN, Z. M. Structural Priming as Implicit Learning: A Comparison of Models of Sentence Production. *Journal of Psycholinguistic Research*, 29(2), 217-229, 2000.

CHOMSKY, N. *Aspects of the theory of syntax.* Oxford: MIT Press, 1965.

DJIKSTRA, A.; VAN HEUVEN, W. J. B. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 23, 175-197, 2002.

DU BOIS, J. W.; CHAFE, W. L.; MEYER, C.; THOMPSON, S. A.; Englebretson, R.; Martey, N. *Santa Barbara corpus of spoken American English, Parts 1-4*. Philadelphia: Linguistic Data Consortium, 2000-2005.

DUARTE, Y. As passivas no português e no inglês: uma análise funcional. *D.E.L.T.A.*, 6(2), 139-167, 1990.

DUSSIAS, P. E. Syntactic ambiguity resolution in L2 learners: Some effects of bilingualism on L1 and L2 processing strategies. *Studies in Second Language Acquisition*, 25, 529-557, 2003.

DUSSIAS, P. E.; SEGARRA, N. The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism, Language and Cognition*, 10(1), p. 101-116, 2007.

ELLIS, N. C. Constructions, Chunking and Connectionism: The Emergence of Second Language Structure. In: DOWTY, C. J.; LONG, M. H. (Ed.). *The Handbook of Second Language Acquisition. Malden*. MA: Blackwell, 63-103, 2003.

GERKEN, L. Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition*, 98, B67-B74, 2006.

GLEITMAN, L.; JANUARY, D.; NAPPA, R.; TRUESWELL, J. C. On the give and take between apprehension and utterance formulation. *Journal of Memory and Language*, 57(4), 544-569, 2007.

GOLLAN, T. H.; SLATTERY, T. J.; GOLDENBERG, D.; RAYNER, K.; ASSCHE, E. V.; DUYCK, W. Frequency drives lexical access in reading but not in speaking: the frequency-lag hypothesis. *Journal of Experimental Psychology: General*, 140(2), 186-209, 2011.

GOLDBERG, A. E. *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago Press, 1995.

GREEN, D. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67-81, 1998.

GRIES, S. T. *Quantitative Corpus Linguistics with R:* A Practical Introduction. New York and London: Routledge, 2009.

GROSJEAN, F. Neurolinguistis, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3-15, 1989.

GROTHENDIECK, G. gsubfn package. Available at https://cran.r-project.org/web/packages/gsubfn/gsubfn.pdf, accessed on August 22nd, 2018.

GUIMARÃES, M. P. A análise da influência translinguística entre o PB e o inglês através da construção passiva. 79 p. *Unpublished master thesis* – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Federal de Minas Gerais, Belo Horizonte. 2016.

GUIMARÃES, M. P.; SOUZA, R. A. Divergências entre a construção passiva no português brasileiro e no inglês: evidências de corpus oral. *Scripta*, 20(38), 262-286, 2016.

HARTSUIKER, R. J.; BERNOLET, S.; SCHOONBAERT, S.; SPEYBROECK, S.; VENDERELST, D. Structural priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, 58, 214-238, 2008.

HARTSUIKER, R. J.; PICKERING, M. J.; VELTKAMP, E. Is syntax separate or shared between languages? Cross-linguistic structural priming in Spanish/English bilinguals. *Psychological Science*, 15, 409-414, 2004.

HERMANS, D.; BONGAERTS, T.; DE BOT, K.; SCHREUDER, R. Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1(3), 213-229, 1998.

HULSTIJN, J. H. *Language Proficiency in Native and Non-native Speakers: Theory and research.* Amsterdam/Philadelphia: John Benjamins Publishing Company, 2015.

HYMES, D. On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics*. Harmondsworth, UK: Penguin Books, 269-293, 1972.

JAEGER, T. F.; SNIDER, N. Implicit Learning and Syntactic Persistence: Surprisal and Cumulativity. In: WOLTER, L.; THORSON, J. (Eds.). *University of Rochester Working Papers in the Language Sciences*, 3(1), 26-44, 2007.

JAEGER, T. F.; SNIDER, N. Alignment as a consequence of expectation adaptation: Structural priming is affected by the prime's prediction error given both prior and recent experience. *Cognition,* 127, 57-83, 2013.

KAHNEMAN, D. *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall, 1973.

KELLO, C. T.; PLAUT, D. C.; MACWHINNEY, B. The task dependence of staged versus cascaded processing: An empirical and computational study of Stroop interference in speech production. *Journal of Experimental Psychology: General*, 129, 340-360, 2000.

KRAMER, R. O efeito de priming sintático na leitura de sentenças na voz passiva por bons e maus leitores dos 5o e 6º anos do ensino fundamental. *Unpublished PhD dissertation*, Programa de Pós-Graduação em Letras – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2017.

KREINER, H.; DEGANI, T. Tip-of-the-tongue in a second language: The effects of brief first-language exposure and long-term use. *Cognition*, 137, 106-114, 2015.

KROLL, J. F.; BOBB, S. C.; WODNIECKA, Z. Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9(2), 119-135, 2006.

KUZNETSOVA, A.; BROCKHOFF, P. B.; CHRISTENSEN, R. H. B. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1–26, 2017.

LA HEIJ, W. Selection processes in monolingual and bilingual lexical access. In: KROLL, J. F.; DE GROOT, A. M. B. (Eds.). *Handbook of bilingualism: Psycholinguistic approaches.* New York: Oxford University Press, 2005, p 289-307.

LEVELT, W. J. M.; ROELOFS, A.; MEYER, A. S. A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1(75), 1999.

MALHOTRA, G.; PICKERING, M. J.; BRANIGAN, H., & BEDNAR, J. A. On the persistence of structural priming: Mechanisms of decay and influence of word-forms. In LOVE, B. C.; MCRAE, K.; SLOUTSKY, V. M. (Eds.). Proceedings of the 30th annual conference of the cognitive science society, 657–662. Austin: Cognitive Science Society, 2008.

MARCUS, M. P., SANTORINI, B., MARCINKIEWICZ, M. A., TAYLOR, A., 1999. Treebank-3.

MATTYS, S. L.; JUSCZYK, P. W.; LUCE, P. A.; MORGAN, J. L. Phonotactic and Prosodic Effects on Word Segmentation in Infants. *Cognitive Psychology*, 38(4), 465-494, 1999.

MAYE, J.; WEISS, D. J.; ASLIN, R. N. Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, 11(1), 122–134, 2008.

MELINGER, A.; DOBEL, C. Lexically-driven structural priming. *Cognition*, 98, B11–B20, 2005.

NATION, I. P. *Teaching and learning vocabulary*. Boston: Heinle & Heinle, 1990.

ORTEGA, L. *Understanding Second Language Acquisition.* London and New York: Routledge, 2009.

PICKERING, M. J.; BRANIGAN, H. P. The Representation of Verbs: Evidence from Structural priming in Language Production. Journal of Memory and Language, 39, 633-651, 1998.

PICKERING, M. J.; FERREIRA, V. S. Structural Priming: A Critical Review. *Psychological Bulletin*, 134(3), 427-459, 2008.

R CORE TEAM. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2013. URL: http://www.R-project.org/.

RASO, T. Artigos Fala e escrita: meio, canal, consequências pragmáticas e linguísticas. *Domínios de Lingu@gem*, 7(2), 12–46, 2013.

RASO, T.; MELLO, H. C-Oral-Brasil I. Belo Horizonte: Editora UFMG, 2012.

RASO, T.; MELLO, H. (Eds.). *Spoken Corpora and Linguistic Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2014.

REEDER, P. A., NEWPORT, E. L, & ASLIN, R. N. From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cognitive Psychology*, 66, 30–54, 2013.

SAFFRAN, J. R.; ASLIN, R. N.; NEWPORT, E. L. Statistical Learning by 8-Month-Old Infants. *Science*, 274(5294), 1926–1928, 1996.

SCHNEIDER, W.; SHIFFRIN, R. M. Controlled and Automatic Human Informaton Processing: I. Detection, Search, and Attention. *Psychological Review*, 84(1), 1-66, 1977.

SEGALOWITZ, N.; HULSTIJN, J. H. Automaticity in Bilingualism and Second Language Learning. In: KROLL, J. F.; DE GROOT, A. M. B. (Eds.). *Handbook of Bilingualism: Psycholinguistic Approaches.* New York: Oxford University Press, 2005, p. 371-388.

SOARES-SILVA, J. Exploring a vocabulary test and a judgment task as diagnoses of early and late bilinguals' L2 proficiency. *Unpublished Ph.D. dissertation* – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Federal de Minas Gerais, Belo Horizonte. 2016.

SOUZA, R. A.; OLIVEIRA, C. S. F. The learnability of the resultative construction in English L2: A comparative study of two forms of the acceptability judgment task. *Revista da Abralin*, 13, 375-410, 2014.

SOUZA, R. A; OLIVEIRA, C. S.; GUIMARÃES, M. P.; ALMEIDA, L. R. Efeitos do bilinguismo sobre a L1: Evidências em julgamentos de aceitabilidade e no processamento online de bilíngues em imersão na L2 ou não. *Revista Linguística,* 10(1), 193-212, 2014.

SOUZA, R. A; OLIVEIRA, C. S. Are Bilingualism Effects on the L1 Byproducts of Implicit Processes? Evidence from Two Experimental Tasks. *Revista de Estudos da Linguagem*, 25(3), 1685-1716, 2017.

TEIXEIRA, M. T. O efeito de priming sintático no processamento de sentenças ativas e passivas no português brasileiro. *Unpublished master thesis.* Pontifícia Universidade Católica do Rio Grande do Sul.

TRUJILLO, A. E. G. Spanish-English Bilinguals' Processing of Two Types of Causative Constructions. *Unpublished master thesis.* Programa de Pós-Graduação em Estudos Linguísticos – Universidade Federal de Minas Gerais, Belo Horizonte, 2018.

ULLMAN, M. T. Contributions of memory circuits to language: the declarative/procedural model. *Cognition,* 92, 231-270, 2004.

VIGLIOCCO, G.; ANTONINI, T.; GARRET, M. F. Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8(4), 314-317.

5. APPENDIX 1: verb passive biases from C-Oral-Brasil I (Raso and Mello, 2012)

verb lemma	passive bias
abaixar	0,0000
abrir	0,0111
acabar	0,0128
aceitar	0,0000
acelerar	0,0000
acender	0,0000
achar	0,0000
acordar	0,0000
acostumar	0,0000
acreditar	0,0000
adiantar	0,0000
adivinhar	0,0000
adorar	0,0000
agüentar	0,0000
ajudar	0,0000
almoçar	0,0000
apagar	0,0000
apertar	0,0000
aprender	0,0000
apresentar	0,0000
aproveitar	0,1429
arrumar	0,0333
assistir	0,0000
atender	0,0909
aumentar	0,0000
beber	0,0000
botar	0,0000
buscar	0,0000
cangar	0,0000
cantar	0,0000
casar	0,0000
chamar	0,0000

verb lemma	passive bias
cobrar	0,0000
colocar	0,0000
combinar	0,0769
começar	0,0000
comer	0,0000
comprar	0,0128
congelar	0,1538
conhecer	0,0135
conseguir	0,0000
consertar	0,0000
contar	0,0000
continuar	0,0000
conversar	0,0000
copiar	0,0000
cortar	0,0800
criar	0,1000
cuidar	0,0000
dar	0,0050
deitar	0,0000
deixar	0,0000
descartar	0,0000
descer	0,0000
descobrir	0,0000
desculpar	0,0000
dever	0,0000
diminuir	0,0000
dividir	0,0556
dizer	0,0000
encher	0,0000
encontrar	0,0000
enfiar	0,0000
enrolar	0,0000

verb lemma	passive bias
ensinar	0,0000
entender	0,0000
entregar	0,0000
errar	0,0000
escolher	0,0000
escrever	0,2059
escutar	0,0000
esperar	0,0000
esquecer	0,0000
estudar	0,0000
explicar	0,0000
falar	0,0000
fazer	0,0191
fechar	0,0000
filmar	0,2000
flagrar	0,0000
foder	0,0833
formar	0,0000
ganhar	0,0274
gastar	0,0000
gravar	0,0588
guardar	0,0217
imaginar	0,0000
jantar	0,0000
jogar	0,0261
juntar	0,0000
lavar	0,0000
lembrar	0,0000
ler	0,0000
levantar	0,0000
levar	0,0066
mandar	0,0000

verb lemma	passive bias
manter	0,0000
marcar	0,0606
matar	0,0000
melhorar	0,0526
meter	0,0000
mexer	0,0000
mostrar	0,0000
mudar	0,0000
olhar	0,0000
operar	0,0000
ouvir	0,0000
pagar	0,0423
parar	0,0000
partir	0,0000
passar	0,0000
pedir	0,0000
pegar	0,0000
pensar	0,0000
perceber	0,0000
perder	0,0816
perguntar	0,0200
picar	0,0000
pintar	0,0870
pôr	0,0082
precisar	0,0000
prender	0,5385
prestar	0,0000
procurar	0,0435
produzir	0,2143
puxar	0,0000
quebrar	0,0000
queimar	0,1000

verb lemma	passive bias
querer	0,0000
receber	0,0000
reclamar	0,0000
resolver	0,0000
rolar	0,0000
roubar	0,0952
salvar	0,0000
segmentar	0,0000
segurar	0,0000
sentar	0,0000
sentir	0,0000
separar	0,1429
servir	0,0000
soltar	0,0455
subir	0,0000
tentar	0,0000
terminar	0,0000
tirar	0,0067
tocar	0,0909
tomar	0,0000
trabalhar	0,0000
tratar	0,1538
trazer	0,0000
trocar	0,0145
usar	0,0448
vender	0,0405
ver	0,0000
virar	0,0278
viver	0,0000
voar	0,0000
zoar	0,0089

6. APPENDIX 2: verb passive biases from SBCSAE (Du Bois et al., 2000-2005)

verb lemma	passive bias
accept	0,000
add	0,000
agree	0,375
answer	0,143
ask	0,022
believe	0,000
blame	0,000
break	0,250
bring	0,097
build	0,327
buy	0,009
call	0,240
carry	0,048
catch	0,250
change	0,040
check	0,000
create	0,059
cut	0,185
dance	0,000
deal	0,000
describe	0,100
do	0,053
draw	0,154
drink	0,000
drive	0,000
drop	0,000
eat	0,000
enjoy	0,000
feed	0,045
feel	0,000
figure	0,063

verb lemma	passive bias
fill	0,115
find	0,023
finish	0,300
follow	0,143
forget	0,000
get	0,000
give	0,095
grab	0,000
hate	0,000
hear	0,000
help	0,000
hit	0,000
hold	0,086
hurt	0,158
include	0,200
involve	0,765
keep	0,014
kill	0,000
know	0,016
learn	0,000
leave	0,098
let	0,000
like	0,000
listen	0,000
look	0,000
lose	0,000
love	0,035
make	0,047
measure	0,200
meet	0,000
mention	0,167

verb lemma	passive bias
miss	0,000
move	0,111
name	0,789
need	0,000
notice	0,071
open	0,085
paint	0,417
pass	0,059
pay	0,175
pick	0,000
play	0,020
pour	0,150
prove	0,167
pull	0,023
push	0,000
put	0,050
raise	0,400
reach	0,000
read	0,012
realize	0,000
receive	0,100
record	0,083
remember	0,000
replace	0,125
run	0,000
say	0,005
see	0,000
sell	0,088
send	0,069
set	0,176
share	0,000

verb lemma	passive bias
shoot	0,000
show	0,114
speak	0,000
spend	0,071
start	0,059
steal	0,000
stick	0,211
take	0,018
talk	0,000
teach	0,067
tell	0,007
think	0,000
throw	0,019
touch	0,083
train	0,300
turn	0,043
understand	0,000
use	0,054
want	0,000
watch	0,020
wear	0,000
win	0,000
work	0,079
write	0,172