

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

Bruna Rodrigues Fontoura

A Study of English L2 Users' Processing Inflectional Morphemes and
their Working Memory Capacity

Belo Horizonte
2018

Bruna Rodrigues Fontoura

A Study of English L2 Users' Processing Inflectional Morphemes and
their Working Memory Capacity

Dissertação apresentada ao Programa de Pós-Graduação
em Estudos Linguísticos da Faculdade de Letras da
Universidade Federal de Minas Gerais, como requisito
parcial para obtenção do título de Mestre em Linguística
Teórica e Descritiva.

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

Belo Horizonte
Faculdade de Letras da UFMG
2018

Ficha catalográfica elaborada pelos Bibliotecários da Biblioteca FALE/UFMG

F684s

Fontoura, Bruna Rodrigues.

A Study of English L2 Users' Processing Inflectional Morphemes and their Working Memory Capacity [manuscrito] / Bruna Rodrigues Fontoura. – 2018.

78 f., enc. : il., tabs., grafs., p&b.

Orientador: Ricardo Augusto de Souza.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: Processamento da Linguagem.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.

Bibliografia: f. 67-72.

Anexos: f. 73-78.

1. Língua inglesa – Morfologia – Teses. 2. Língua inglesa – Flexão – Teses. 3. Aquisição da segunda linguagem – Teses. 4. Língua inglesa – Falantes de português – Brasil – Teses. I. Souza, Ricardo Augusto de. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD : 418.02



UNIVERSIDADE FEDERAL DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGÜÍSTICOS



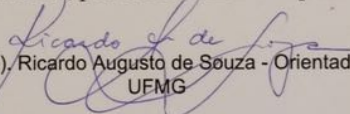
FOLHA DE APROVAÇÃO

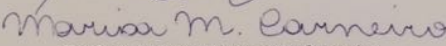
**A Study of English L2 Users' Processing Inflectional Morphemes
and their Working Memory Capacity**

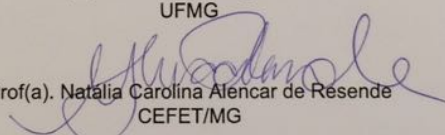
BRUNA RODRIGUES FONTOURA

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

Aprovada em 16 de fevereiro de 2018, pela banca constituída pelos membros:


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


Prof(a). Marisa Mendonça Carneiro
UFMG


Prof(a). Natália Carolina Alencar de Resende
CEFET/MG

Belo Horizonte, 16 de fevereiro de 2018.

To my beloved and caring family.

ACKNOWLEDGEMENTS

First of all, I would like to thank Prof. Dr. Ricardo de Souza for accepting to advise me. Thanks for the patience and good insights on this study.

Thanks to Prof. Dr. Rui Rothe-Neves for taking same time to talk to me and for the wonderful suggestions for this study.

Thanks to Dr. Michael Kane who so gently and promptly sent me the n-back tasks.

Thanks to Prof. Dr. Marisa Carneiro for being part of this process.

Thanks to all the professors from IngRede, I am sure I feel much more critical and qualified because of you.

Thanks to all the amazing professors I have had during this process, you made me question and rethink a lot.

Thanks to Prof. Dr. Marisa Carneiro and Prof. Dr. Natalia Resende for all the contributions to this study.

Special thanks to my parents, Flávio and Márcia, for supporting me and being there for me all the time. I know we have been living in different cities for almost eight years, but it seems as if I have never left home. Thanks for always believing in me and letting me follow my dreams.

I would like to thank my grandparents, Odracy and Antônio Carlos, and my aunt, Flávia, who have always been so proud of me. I know that wherever you are, grandpa, you are still very proud of me. This is our victory!

Thanks to my dear cats, Elis and Nina, for showing me that love is so simple and so precious. I will always miss Elis living with me, but I can accept the idea of visiting her on the holidays.

Special thanks to my dearest Fábio for being with me during this process, nothing of this would have been possible without you. Thanks for helping me out with programming, statistics, and always make me go even further. I am sure I am a better person because of you!

I would like to thank my friends. I know that putting up with all my whining is not easy, your advice and support was so important during this process.

Thanks to everybody from the Psycholinguistics Lab at UFMG who advised me and helped me in so many ways, you are all amazing.

Thanks to all the people who willingly contributed to this study.

Thanks to CAPES for the scholarship.

Thanks to everyone who has been part of this!

“Whatever you are, be a good one.”
(Anonymous author)

ABSTRACT

Inflectional morphology poses many challenges for second language (L2) learners of English, as these individuals face problems in processing. Some researchers, such as Jiang (2004b, 2007) and Carneiro (2011), have investigated how L2 English learners process inflectional morphology. They have found evidence showing that L2 learners were not sensitive to violations of inflectional morphemes. In addition, McDonald (2006) claims that L2 processing difficulty is related to some factors among which working memory capacity plays a significant role. Based on that, this study aims to investigate if working memory capacity influences English L2 learners processing ability of the English inflectional morphemes: third person singular (*-s*) agreement, and past regular verbs (*-ed*). In order to do so, participants were expected to conduct three experiments in person and a proficiency test online. First, two off-line experiments in the format of acceptability judgment tasks, one with a time constraint and another with a memory load – both containing grammatical and ungrammatical sentences –, and a 2-back task, in which similar letters had to be selected from a list, were carried out by participants. The software PsychoPy was used to present the stimuli, and record the reaction time and performance of participants. Proficiency was measured with the Vocabulary Levels Test (VLT). Participants comprised a group of Brazilians native speakers of Portuguese and English as an L2 that were divided into high and low proficiency. In the acceptability judgment tasks, the reaction time and score ranking of the high proficient group were compared concerning their performance on grammatical and ungrammatical sentences and the same was done for the low proficient group. The reaction time and performance of high proficient participants were compared to the low proficient participants in the 2-back task. In what concerns the acceptability judgment tasks, the results indicate that the time constraint and working memory load did not offer an onus to processing of the inflectional morphemes under investigation. This suggests that working memory does not influence morphological processing. Moreover, when the performance of the high and low proficient participants was compared in the 2-back task, there was no statistical significance, which indicates that level of proficiency does not affect working memory capacity. The results obtained are discussed under the Relational Morphology proposed by Jackendoff & Audring (2016).

Keywords: second language; processing; inflectional morphology; working memory capacity

RESUMO

A morfologia flexional apresenta muitos desafios para aprendizes de inglês como segunda língua (L2), já que esses indivíduos enfrentam problemas no processamento. Alguns pesquisadores, tais como Jiang (2004b, 2007) e Carneiro (2011), investigaram como aprendizes de inglês como L2 processam a morfologia flexional. Eles encontraram evidências que mostram que os aprendizes de L2 não eram sensíveis à violações de morfemas flexionais. Além disso, McDonald (2006) declara que o processamento de L2 é difícil devido a alguns fatores, entre os quais a capacidade de memória de trabalho desempenha um papel significativo. Com base nisso, este estudo busca investigar se a capacidade de memória de trabalho influencia a habilidade de processamento de aprendizes de inglês como L2 no que concerne aos morfemas flexionais da língua inglesa: terceira pessoa singular (-s) e verbos regulares passados (-ed). Para tanto, os participantes realizaram três experimentos pessoalmente e um teste de proficiência *online*. Em primeiro lugar, dois experimentos *off-line* no formato de tarefas de julgamento de aceitabilidade, uma com restrição temporal e outra com carga de memória de trabalho – ambas contendo frases gramaticais e agramaticais –, e uma tarefa de *2-back*, na qual letras similares deveriam ser selecionadas de uma lista, foram realizados pelos participantes. O software *PsychoPy* foi usado para apresentar os estímulos, e registrar o tempo de reação e a performance dos participantes. A proficiência foi medida com o *Vocabulary Levels Test* (VLT). Os participantes consistiram em um grupo de brasileiros falantes nativos de português e de inglês como L2 que foram divididos em alto e baixo nível de proficiência. Nas tarefas de julgamento de aceitabilidade, o tempo de reação e a classificação feita pelo grupo de alta proficiência foram comparadas ao seu desempenho em frases gramaticais e agramaticais, e o mesmo foi feito para o grupo de baixa proficiência. O tempo de reação e o desempenho dos participantes de alta proficiência foram comparados aos dos participantes de baixa proficiência na tarefa de *2-back*. No que diz respeito às tarefas de julgamento de aceitabilidade, os resultados indicam que a restrição temporal e a carga de memória de trabalho não ofereceram um ônus ao processamento de morfemas flexionais sob investigação. Isso sugere que a memória de trabalho não influencia o processamento morfológico. Além disso, quando o desempenho dos participantes de alta e baixa proficiência foi comparado na tarefa de *2-back*, não houve significância estatística, o que indica que o nível de proficiência não afetou a capacidade de memória de trabalho. Os resultados obtidos são discutidos com a Morfologia Relacional proposta por Jackendoff & Audring (2016).

Palavras-chave: segunda língua; processamento; morfologia flexional; capacidade de memória de trabalho

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION.....	15
1.1 Preliminaries	15
1.2 Statement of Purpose	16
1.3 Research Objectives	17
1.3.1 General Objectives	17
1.3.2 Specific Objectives	18
1.4 Dissertation Organization	18

CHAPTER 2

THEORETICAL BACKGROUND	19
2.1 Relational Morphology (RM)	19
2.2 Declarative/Procedural Model and the Second Language Acquisition.....	27
2.3 Inflectional Morphemes studies.....	29
2.4 Memory Studies	32

CHAPTER 3

METHODOLOGY	35
3.1 Overview	35
3.2 Experiment 2: Acceptability Judgment with Time Restriction	36
3.3 Experiment 3: Acceptability Judgment with Memory Load	37
3.4 Experiment 4: 2-back	38
3.5.1 Materials	39
3.5.2 Acceptability judgment (AJ) task	39
3.5.2.1 Acceptability judgment task with time restriction (AJT)	40
3.5.2.2 Acceptability judgment task with memory load (AJWM).....	42
3.5.3 N-back task	45
3.5.4 Vocabulary Levels Test (VLT).....	46

CHAPTER 4

DATA ANALYSIS	49
4.1 Statistical analysis	49
4.2 Participants	49
4.3 Experiment 1: Acceptability Judgment with Time Restriction	51

4.4 Experiment 2: Acceptability Judgment with Memory Load	53
4.4.1 Acceptability Judgment	53
4.4.2 Memory Load	56
4.5 Experiment 3: 2-back	58
CHAPTER 5	
CONCLUSION	62
5.1 Discussion.....	63
5.2 Limitations of the Study.....	65
REFERENCES	67

LIST OF APPENDIXES

APPENDIX 1- Vocabulary Levels Test (VLT)	73
APPENDIX 1 – Acceptability Judgment with Time Restriction	74
APPENDIX 2 – Acceptability Judgment with Memory Load	76
APPENDIX 3 – 2-back	78

LIST OF TABLES

Table 1. Structure of the research	36
Table 2. Acceptability Judgment Task Likert Scale (SILVA, 2016, p. 61)	41
Table 3: Means and SD of Participants' RT in the AJT task	47
Table 4: Shapiro-Wilk test for Participants' RT distribution in AJT	51
Table 5: Means and Medians of Participants' sentence score in the AJT task.....	52
Table 6: Shapiro-Wilk test for Participants' sentence score distribution in AJT	52
Table 7: Means and SD of Participants' RT in the AJWM task	53
Table 8: Shapiro-Wilk test for participants' RT distribution in AJWM	54
Table 9: Means and Medians of Participants' sentence score in the AJWM task	54
Table 10: Shapiro-Wilk test for participants' sentence score distribution in AJWM	55
Table 11: Shapiro-Wilk test for participants' Levenshtein distance distribution in AJWM...57	57
Table 12: Means and SD of participants' RT in the 2-back task	58
Table 13: Shapiro-Wilk test for participants' RT distribution in 2-back	59
Table 14: Means and SD of Participants' performance in the 2-back task	60
Table 15: Shapiro-Wilk test distribution for participants' performance in 2-back	60

LIST OF FIGURES

Figure 1: Baddeley’s (1992b) working memory model p. 557	32
Figure 2: Matching in VLT (SOUZA & SILVA, 2015, p. 193)	47
Figure 3: Participants’ ages	50
Figure 4: Grammatical and Ungrammatical Sentences Score Means of High and Low Proficient Participants in AJT	52
Figure 5: Grammatical and Ungrammatical Sentences Score Means of High and Low Proficient Participants in AJWM	54
Figure 6: High and Low Proficient Participants Means of Levenshtein Distance in AJWM	56
Figure 7: ROC curve of High and Low participants’ performance in Memory Load in AJWM	57
Figure 8: ROC curve of High and Low participants’ RT in 2-back	59
Figure 9: ROC curve of High and Low participants’ performance in 2-back	60

CHAPTER 1

INTRODUCTION

1.1 Preliminaries

This research addresses the issue of L2 inflectional morphemes processing considering the way Brazilian learners of English as a Second Language (ESL) process sentences. It is also taken into consideration if ESL learners' proficiency level and working memory capacity influence their ability to process inflectional morphemes.

“The ‘procedures’ that comprise implicit knowledge can be easily and rapidly accessed in unplanned language use. In contrast, explicit knowledge exists as declarative facts that can only be accessed through the application of attentional processes” (ELLIS *et al.*, 2009 p. 12). Therefore, the users of a language know, for example, when they are supposed to add *-ed* to the end of the verbs because they are talking about something that happened in the past, that is, this knowledge is automatic. However, learners of a second language (L2) that have not automatized their L2 yet make use of explicit knowledge to monitor their language. This is the reason why they may make mistakes in on-line processing where there is little time to think and use declarative information.¹

It is argued by Sorace (2006b, 2011) that near-native speakers have native-like representations as well as syntactic knowledge, considering the syntax-pragmatics interface. The syntax-pragmatics interface encompasses formal features of the language system. Furthermore, Sorace (2006b, 2011) defends that any difficulty presented at this advanced level may be associated with access and integration of syntactic knowledge. As a result, Sorace (2011) claims that there is no substantial indication that the Shallow Processing hypothesis defended by Clahsen & Felser (2006a, 2006b) applies to near-native like speakers, who have acquired such features. Clahsen & Felser (2006a) defend that “the syntactic representations adult L2 learners compute for comprehension are shallower and less detailed than those of native speakers” (p. 32). Therefore, while lexical-semantic information is similar to native speakers in parsing, syntactic information is limited.

As pointed out by Jiang (2004b, 2007) and Carneiro (2011), ESL learners are not sensitive to inflectional morphemes violations – such as regular plural (-s) and third person singular (-s) agreement – in on-line experiments using the self-paced reading with the

¹ It is important to point out here that this view is shared by many authors, such as Ellis *et al.* (2009); however, Jackendoff & Audring (2016) claim that rules are part of the declarative knowledge of the speaker, and therefore, are part of the explicit knowledge. In this study, we share Jackendoff & Audring's (2016) view.

moving-window paradigm. In both Jiang's studies (2004b, 2007), native speakers noticed ungrammatical constructions; nonetheless, Carneiro's study (2011) showed that native speakers did not exhibit significant differences compared to the ESL learners in relation to third person singular (-s) violations while they were sensitive to the regular past (-ed) violations:

(1) The smiling waiter/ always/ serve(s)/ the clients/ at the French restaurant. *Third Person Singular (-s) agreement* (CARNEIRO, 2011, p. 144)

(2) After/ the fire/ stopped/ completely/ the firefighter/ help(ed)/ the victims/ with severe burns. *Past Regular Verbs (-ed)* (CARNEIRO, 2011, p. 133)

(3) They discovered one of the dinosaur(s) was different from the others. *Regular plural (-s)* (JIANG, 2007, p. 30)

Marinis (2010) argues that in off-line comprehension tasks, individuals that have a fairly high working memory capacity may have a better performance than those with a fairly low working memory. He also proposes that on-line comprehension tasks require fewer resources from working memory.

According to Gathercole (2007), working memory plays a significant role in language processing regarding competent language users and during the development of a language as well. Cowan (2005, 2010) claims that working memory is a key factor in mental tasks, for instance, language comprehension, since ideas from previous sentences may be combined with other ones.

Accordingly, this study evaluated whether Brazilian learners of English are sensitive to inflectional morphemes, such as third person singular (-s) and regular past verbs (-ed). In order to do so, two off-line experiments were conducted, an acceptability judgment (AJ) task with two versions: a speeded acceptability judgment task similar to the one in Souza & Silva (2015) and Silva (2016), and another one with external memory load, adapted from McDonald (2006, 2008b). A working memory test was carried out based on the study of Kane *et al.* (2007) about n-back. The proficiency level of the participants was measured according to the studies conducted by Souza & Silva (2015), Souza; Duarte & Berg (2015) and Silva (2016) regarding the Vocabulary Levels Test (VLT). Additionally, the Relational Morphology (RM) proposed by Jackendoff & Audring (2016) was used as a theoretical background to corroborate the investigation that was performed.

1.2 Statement of Purpose

Slabakova (2014) argues that functional morphology is “the sticking point of acquisition” since it postulates grammar formal features. According to her, functional morphology is not only a challenge for L2 learners, but also for children and adult native speakers who are not attentive to syntactic features. Hence, functional morphology is difficult when it comes to language production and comprehension. It demands “a higher cognitive load in processing because it carries higher syntactic information.” (SLABAKOVA, 2014, p. 13)

While acquisition of a child’s first language (L1) becomes automatic, adult learners of a second language (L2) face more difficulty in acquiring some grammatical features than others, once some aspects of L2 have fossilized (JIANG, 2007). According to Han (2013, p. 159), “acquisition of grammatical morphemes hinges on conceptual restructuring, something that does not happen easily, if at all”. Han (2010, 2013) argues that despite the fact that the L2 learners have explicit knowledge about how plural (-s) must be used, they cannot properly employ it in normal conversation. Although the present study did not focus on reading comprehension such as Han (2010, 2013), this study investigated how L2 English learners process inflectional morphemes in acceptability judgment tasks, as Slabakova (2014) defends that functional morphology is the bottleneck of second language acquisition. Besides, Jiang (2004b, 2007), McDonald (2006, 2008b) and Carneiro (2011) have carried out studies where sentence comprehension of inflectional morphemes was under investigation².

“Young adults can recall three to five chunks from a presented list no matter whether these are learned pairs or singletons” (COWAN, 2010 p. 52). When a person is reading, the individual can only proceed if all the items have been integrated into a whole unit; otherwise it is not possible to continue the operation. As a consequence, we believe that an acceptability judgment task with an external memory load in which a random seven-digit number was presented to participants, similarly to the experiment that McDonald (2006, 2008b) carried out, could attempt to investigate if individuals’ working memory affected their processing ability of inflectional morphemes. Moreover, a speeded AJ was conducted in order to check the differences in the experiment paradigm.

1.3 Research Objectives

1.3.1 General Objectives

² Despite the fact that Jiang (2004b, 2007) and Carneiro (2011) investigated sentence comprehension with on-line experiments – self-paced reading –, this study decided to follow McDonald (2006, 2008b) and adopt off-line experiments since Marinis (2010) pointed out that off-line experiments access more resources from working memory than on-line experiments.

This study aims to delve into whether Brazilian learners of ESL present sensibility in processing the inflectional morphemes –s and –ed in grammatical and ungrammatical sentences. It also investigated L2 learners' proficiency level to check if proficiency is accountable for processing efficiency. Finally, participant's working memory capacity (WMC) is studied in order to see whether individuals with higher working memory capacity also have higher processing capacity.

1.3.2 Specific Objectives

- a) Check if proficiency influences reaction time (RT) and sentence score in AJ tasks when performance of high and low proficient participants are compared in grammatical and ungrammatical sentences within each group.
- b) Investigate if participants display significant differences in their RT and sentence score in the two judgments of acceptability – the one with temporal restriction and the one with external memory load.
- c) Explore whether working memory plays a significant role in processing capacity.
- d) Compare the performance of high and low proficient participants in the AJ with external memory load to the 2-back task to observe if the results are a consequence of proficiency or WMC.

1.4 Dissertation Organization

This dissertation is divided into five distinct chapters, including the introduction. In Chapter 2, we present the theoretical background that we used to explain the results found in this study. Chapter 3 concerns the methods and materials used to develop this research. Chapter 4 deals with the data analysis and interpretation of the results. In Chapter 5, we discuss our findings and limitations of this study.

CHAPTER 2

THEORETICAL BACKGROUND

2.1 Relational Morphology (RM)

The Relational Morphology (RM) proposed by Jackendoff & Audring (2016) seeks to integrate phonology, morphology, syntax and semantics. They believe “a theory of the language faculty should seek integration with theories of other mental faculties. However, such integration must respect the details of the individual faculties” (p. 468). The authors base their model on the studies conducted by Jackendoff (2002) and Culicover & Jackendoff (2005) regarding the Parallel Architecture and also on Booij’s (2010) Construction Morphology. Their aim is to delve into “some aspects of a theory of morphological processing” (p. 468).

They follow the approach defended by proponents of the Parallel Architecture, Construction Morphology, and Construction Grammar (GOLDBERG, 1995³ apud JACKENDOFF & AUDRING, 2016, p. 468; HOFFMAN & TROUSDALE, 2013⁴ apud JACKENDOFF & AUDRING, 2016, p. 468), according to which “grammatical rules are themselves lexical items – that is, the grammar is part of the lexicon.” (p. 469). Below we find the explanation that supports their reasoning:

in practically every linguistic theory, a word contains pieces of structure on three levels: its semantic structure, its syntactic features, and its phonology. In the Parallel Architecture framework, these levels are in principle independent, each with its own characteristic conditions of well-formedness. But each is also linked to the others: *this* phonological string can be linked to *this* syntax and *this* semantics. We call these connections interface links, and we notate them with indices that show which parts of structure on one level correspond with structure on another level. [...] The indices are to be thought of as marking the ends of association lines (JACKENDOFF & AUDRING, 2016, p. 469).

(4) Mapping of the morphologically complex word *sheepish* that illustrates the previous explanation (JACKENDOFF & AUDRING, 2016, p. 469):

Semantics:	[SHEEP ₁ -LIKE; TIMID] ₂
Morphosyntax:	[N – aff]
	A 1 3 2
Phonology:	/ʃi:p ₁ iʃ _{3/2} /

³ GOLDBERG, A. *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press, 1995.

⁴ HOFFMAN, T., & TROUSDALE, G. (Eds.). *The Oxford handbook of Construction Grammar*. Oxford: Oxford University Press, 2013. doi: 10.1093/oxfordhb/9780195396683.001.0001

According to Jackendoff & Audring (2016), the morphosyntax determines that this word is an adjective constituted of a noun and an affix. The numbers in (4) correspond to coindexes, this coindexation functions as a relational link. Hence, the syntactic category Noun, the meaning SHEEP and the phonology /ʃip/ are connected by coindex 1. The semantics, morphosyntax, and phonology of the entire word are connected by coindex 2. The affix and its phonology /ɪʃ/ are connected by coindex 3. Nonetheless, the significant chunks of *sheepish* and *sheep* are also connected by coindex 1.

While Construction Grammar takes the example in (5) below as a construction, the authors follow Construction Morphology, and name it a schema. The authors defend that there is a significant difference of exchanging rules by schemas. “Schemas have the same format as words, differing only in that part of their structure consists of variables and variable coindices. Hence words and schemas can be in the same ‘place’ in the theory – or in the mind” (JACKENDOFF & AUDRING, 2016, p. 471). They claim that it is not necessary to have the construct of “lexicon” and “grammar” independently in this theory. Additionally, since schemas can seize what is similar among lexical items, they “take less work to learn, store, and/or process” (JACKENDOFF & AUDRING, 2016, p. 472). This is the relational function of schemas.

As they emphasize, the concept of schema marks that it is a declarative rather than a procedural representation. Jackendoff & Audring (2016) claim that (5) “is not a procedure to convert an “input” into an “output,” by adding *-ish* to a noun to form an adjective; rather, it licenses or motivates its instances” (JACKENDOFF & AUDRING, 2016, p. 471). This schema has three layers of structure, such as (4). Nevertheless, it has open slots or variables rather than a lexical base:

- (5) Representing words that follow the same pattern as in *sheepish* in (4)
(JACKENDOFF & AUDRING, 2016, p. 471)

Semantics: $[X_y\text{-LIKE}]_z$
 Morphosyntax: $[\text{N} - \text{aff}]$
 Phonology: $/\dots y \text{I} \text{ʃ} \text{z} /$

The affix in the morphosyntax and its phonology /ɪʃ/ are connected by the coindex 3. Conversely, the variable coindex *z* connects the layers of the schema completely and the variable coindex *y* connects the variables inside the schema.

Besides, the authors claim that schemas are not only used for lexical items, but also

for idioms, such as *miss the boat*. In the later case, “a schema for the regular past tense can apply to motivate these stored forms. It is the same schema, the same generalization, but instantiated here through relational rather than generative means” (JACKENDOFF & AUDRING, 2016, p. 472). Furthermore, Baayen, Dijkstra, & Schreuder (1997) have presented data showing that some nouns of Dutch are stored instead of being formed generatively, as it saves time to avoid subcategorization conflict. This is a very complicated issue in a rule-based theory as items that can be generated through rules are not supposed to be in the lexicon. According to Langacker (1987), the “rule-list fallacy” happens when:

one is forced to choose between rules and lists: the options are posed as rules alone vs. lists alone. If these are the only two options, it can be argued that the rules must be chosen, for lists by themselves do not express generalizations. There is in reality a third choice, however, namely both rules and lists (p. 42).

Langacker (1987) illustrates this by saying that English speakers not only learn the general rule of plural formation, but also a lot of regular plural forms in the format of fixed units. Moreover, the speakers master – in the format of fixed units – a great quantity of conventional expressions that are completely “analyzable and regular in formation”.

As a consequence, Jackendoff & Audring (2016) argue that there are schemas that present both the generative and the relational functions. Nonetheless, there is no schema with only the generative function, “because any instance of a schema that can be generated online can then be stored [...] One can store items of all sizes [...] without losing the internal structure that connects them to the grammatical patterns in the language” (p. 473). Therefore, there are two types of schemas, productive schemas and nonproductive schemas. The first one performs the function of traditional generative rules and seizes generalizations in stored lexical items, and the latter only encompasses the relational function.

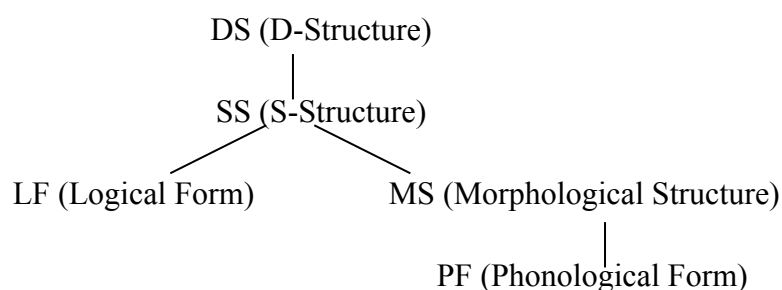
The authors defend that “the relational function is the foundation on which the generative function builds” (p. 474). Hence, it is necessary to notice that rules/schemas operate in a relational function so that it is possible to understand that the relations within the lexicon cannot be explained considering traditional generative rules. According to them, Halle and Marantz (1993) tried to handle nonproductive schemas by postulating them as generative rules in the Distributed Morphology (DM).

While DM accounts for insertion of inflectional morphemes considering a fixed generative component, Relational Morphology defends that the morphological system is stored in the declarative memory. This could explain how L2 learners that acquire such morphological features would be able to process inflectional morphology in real time, as long

as their processing capacity, such as their working memory for instance, allow them to do so.

Besides, the DM proposed by Halle and Marantz (1993) argues that some syntactic processes are not related to the ones that can be found in phonological processes. Hence, DM allows us to map morphological phenomena without making reference to phonology. DM encompasses grammar according to “principles-and-parameters”. As a consequence, the Morphological Structure (MS) is added to the hierarchical arrangement between syntax and phonology. MS is seen as a syntactic representation element that is part of the phonology, and the latter is considered the interpretative component that comprehends syntactic representations phonologically:

(6) Grammar Structure in DM (Halle and Marantz, 1993, p. 114)



Representations at each of the five levels consist of hierarchical groupings of terminal elements graphically represented by the familiar tree diagram. The terminal elements of the trees consist of complexes of grammatical features. These terminal elements are supplied with phonological features only after Vocabulary Insertion at MS (HALLE & MARANTZ, 1993, p. 114).

DM claims that syntax does not control lexical items, instead, it creates structures when it combines and matches morphosyntactic features in many processes, such as Merger and Fusion. Furthermore, Halle and Marantz (1993) criticizes Chomsky (1993) proposal in relation to affixation.

In addition, the Minimalist Program postulates that lexical items are already inflected in the lexicon; that is, the computing system organizes those items, and then checks them. Accordingly, Minimalism uses the X-bar theory to check the formal features of every item. Despite the fact that this theory does not analyze the clauses in relation to Deep Structure (D-Structure), and Surface Structure (S-Structure), the constituents' movement is allowed so that the features may be checked. Consequently, this postulation is only worried about Phonological Form (PF) and Logical Form (LF). When all items are checked, they are sent to the PF (CHOMSKY, 1993, 1995).

Hence, for DM, vocabulary can be considered part of “the knowledge that speakers have about the interrelationship between the morphosyntactic feature bundle characterizing a

morpheme and its phonological features, that is, about the mapping of morphosyntactic features onto complexes of phonological features” (HALLE & MARANTZ, 1993, p. 123). Thereby, we can see an example of vocabulary insertion in the formation of third person singular present [3sg] in English. The default third person singular present morpheme in English is /-z/:

(7) /-z/ \longleftrightarrow [3sg] (HALLE & MARANTZ, 1993, p. 126).

Furthermore, Jackendoff & Audring (2016) explain how a schema “motivates its instances”. The authors declare that “word families like *sheep*, *sheepish* and [A N-*ish*] motivate each other through an inheritance hierarchy” (p. 474):

(8) $\begin{array}{c} \text{sheep} \quad [\text{A N-ish}] \quad \text{Nchild} \\ \diagdown \quad \diagup \quad \diagdown \\ [\text{A Nsheep ish}] \quad [\text{A Nchild ish}] \end{array}$ (JACKENDOFF & AUDRING, 2016, p. 474)

Therefore, the items that are lower in the hierarchy inherit the structure from the ones that are higher linked in the hierarchy. As a consequence, *sheepish* comes from *sheep* and the last one from [A N-*ish*]. According to the authors, “inheritance is a domain-general theoretical construct that requires no special machinery for morphology or even for language per se” (p. 474).

The authors follow Jackendoff’s (1975) concept of impoverished entry theory which states that the information existing in lexical items cannot be inherited from anywhere else. Hence, it is not necessary to specify any information in a lower node that is contained in a higher one. They instantiate this by saying that *sheepish* can be represented as in (9) since it has features from the two higher nodes:

(9) $\begin{array}{c} \text{Nsheep} \quad [\text{A N-ish}] \\ \diagdown \quad \diagup \\ [\dots; \text{TIMID}] \end{array}$ (JACKENDOFF & AUDRING, 2016, p. 475)

Jackendoff & Audring (2016) do not agree with the impoverished entry theory’s construct that defends that redundant details are deleted. They argue that when a person acquires a new complex word, it is necessary to find what are its details, and just after doing so, it is possible to decide what schemas fit to it and which is the base. Nonetheless, the impoverished entry theory claims that once the new word’s connection to a base and to a schema or more than one is established, the features considered redundant are swept away,

which the authors consider unlikely following Langacker (1987) and Booij (2017).

Booij (2017) criticizes the impoverished entry theory by saying that it does not account for the vastness of our lexical memory. As a consequence, he defends the full entry-theory, also present in Jackendoff (1975, 1997). Booij (2017) argues that memorized complex words are fully specified, and therefore, they are responsible for the acquisition of morphological schemas. The author says that a person acquires the English deverbal noun morpheme *-er* – such as in *baker* and *writer* – after enough exposure of words of the same type. However, it is implausible that the person would erase the predictable information about individual complex words once the morpheme is acquired (JACKENDOFF, 1997; HUDSON, 2007⁵ apud BOOIJ, 2017, p. 22).

Jackendoff & Audring (2016) believe that economy is not the most appropriate standard to “a theory of lexical storage in the brain”. Thus, they argue that “the brain embraces redundancy, at least up to a point” (p. 476). The role of redundancy is defended by Libben (2006) who claims that it makes mental computation more robust. As a result, Jackendoff & Audring (2016) decide to adopt a full-entry theory, which defends that “lexical items are encoded in their entirety, even when redundant” (p. 476).

When children are acquiring a language, they need to figure out the productive rules of that language, therefore, the productive schemas, based on linguistic input. “This requires appropriate evidence to be stored in memory, in the form of stored words and fixed phrases” (JACKENDOFF & AUDRING, 2016, p. 478).

Since it is not enough to only store items, it is necessary to have some process that looks for similarities among items in the long-term memory (LTM). Accordingly, items that have a common structure share sister links, such as *sheep* and *sheepish*. Common parts in the linked sister constitute “the constants in the schema” – such as *-ish* – and the different parts are substituted by variables – such as *sheep*, *child*, *slug*. “By invoking a schema, the learner does not have to compare the new word unselectively with all potential sisters, whatever their degree and dimensions of similarity. The schema in effect offers a prestructured or “precompiled” basis for analogy” (JACKENDOFF & AUDRING, 2016, p. 479).

The authors claim that knowledge of language is a network that has linked nodes. Therefore, the nodes contain three layers of structures that are semantics, morphosyntax and phonology – such as in some versions of lexical networks as in Levelt, Roelofs & Meyer (1999). Jackendoff & Audring (2016) defend that “a “lexical node” is a complex of semantic,

⁵ HUDSON, R. *Language networks: The new word grammar*. Oxford: Oxford University Press, 2007.

morphosyntactic, and phonological nodes connected by interface links” (p. 480). When one considers the relation amid words in lexical networks, there is no simple link among nodes, instead, nodes contain internal structures that are connected by relational links. Hence, the authors claim that the representational theory controls aspects of the processing theory.

The researchers defend that the lexicon enables spreading activation. Therefore, they suggest an adjustment from the representational theory. Having nodes connected by interface links and relational links goes towards the direction of spreading activation. In what concerns the spreading activation, its strength is determined by the extension of the connection.

The construct of spread activation is also discussed in Levelt, Roelofs & Meyer (1999), which can explain how items are selected in production and why word form information is not as crucial as the other aspects. “Lexical selection is conceived of as selecting the syntactic word. What the speaker selects from the mental lexicon is an item that is just sufficiently specified to function in the developing syntax” (LEVELT, ROELOFS & MEYER 1999, p. 14). Thus, the word’s syntax is necessary to produce fluent speech.

Jackendoff & Audring (2016) also make some predictions about processing that are not connected to the representational theory. Based on pure network models, the authors claim that:

working memory (WM) is to be regarded as a functional component distinct from long-term memory (LTM); it does not consist simply of the parts of LTM that are active. LTM contains the lexical network of “knowledge of language” [...] It is in LTM that schemas fulfill their relational function, through their links to more fully specified items. In contrast, working memory (WM) is the functional component in which pieces of lexical items are assembled into larger structures, either to create an utterance (in production), or to analyze and parse an input (in comprehension) (JACKENDOFF & AUDRING, 2016, p. 480, 481).

The authors’ concept of the architecture of memory is in line with the construct of MUC (memory, unification, control), which are components of language processing, as defended by Hagoort (2005):

the Memory component comprises a specification of the different types of language information stored in long-term memory, as well as the retrieval operations. The Unification component refers to the integration of lexically retrieved information into a representation of multi-word utterances. The Control component relates language to action, and is invoked, for instance, when the correct target language has to be selected (in the case of bilingualism), or for handling turn taking during conversation (p. 416).

Since WM is the place where productive schemas are able to carry out their generative function, it is compatible to the concept of “unification” suggested by Hagoort (2005).

Moreover, Jackendoff & Audring (2016) propose that:

processing is “promiscuous”: in language comprehension, everything in LTM sufficiently similar to the current input is activated and retrieved into WM, where it serves as a candidate for “what is being heard,” in competition with other candidates. The degree to which any particular candidate is activated in WM depends on a number of factors, including its current level of activation in LTM [...], the relative strength of its competitors, and how well it fits the current context (p. 481).

The researchers argue that competition between minimal pairs cannot be solved promptly. Therefore, it is necessary to have word identification and syntactic structure beforehand. They exemplify this competition in (10) by explaining that the ambiguity is caused by the words in italic, and it can only be solved after the semantic processing of the words underlined:

(10a) That’s not *a parent*; it’s actually a teacher.

(10b) That’s not *apparent*; it’s actually quite obscure. (JACKENDOFF & AUDRING, 2016, p. 481).

Jackendoff & Audring (2016) also account for the connection of speed of lexical access and corpus frequency of the target word, which is proposed by Oldfield & Wingfield (1965). As a result, there are two correlations involved. The first one involves the correlation between reaction time and resting activation; therefore, faster response corresponds to higher level of resting activation (BAAYEN, DIJKSTRA & SCHREUDER, 1997; BYBEE, 1985⁶, 1995⁷ apud JACKENDOFF & AUDRING, 2016, p. 481; PLAG, 2003⁸ apud JACKENDOFF & AUDRING, 2016, p. 481). Second, the researchers argue that repeated use increases the level of resting activation of an item, as well as common aspects of the Hebbian learning rule⁹. “Thus there is no direct representation of frequency in the brain. Rather, frequency in a written corpus stands proxy for frequency of use, which affects resting activation – which in turn affects reaction time”. (JACKENDOFF & AUDRING, 2016, p. 482). As a result, they conclude that the representational theory controls the structural part and the theory of processing controls the resting activation part.

Jackendoff & Audring (2016) also discuss the activation of a node in the LTM

⁶ BYBEE, J. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins. 1985. doi: 10.1075/tsl.9

⁷ BYBEE, J. Regular morphology and the lexicon. *Language and Cognitive Processes*, v. 10, p. 425–455, 1995. doi:10.1080/01690969508407111

⁸ PLAG, I. *Word-formation in English*. Cambridge: Cambridge University Press, 2003. doi: 10.1017/CBO9780511841323

⁹ The Hebbian Learning Rule states that activation is a key factor in determining the level of connection between two items since it increases or decreases depending on the activation. (For further information see Hebb (1949)).

network, in what concerns the mechanisms to raise the level of activation, they declare that:

after an item is called by WM, we find it reasonable to assume that it does not return to resting activation immediately, but takes a little while to settle down. This provides an account of identity priming: if the item recurs in the input soon enough, its level is still above resting activation, and thus it is easier to summon it back to WM, resulting in more rapid identification (p. 482).

Moreover, the activation of a node can also happen when its activation comes from a relational link of a lexical item activated in LTM, which is also faster to answer WM. This process is called neighboring priming.

The authors argue that schemas are lexical entries, therefore, they also contain resting activation accounted for the WM. Besides, their strength can be measured in relation to their frequency, which is connected to syntactic priming.

In what concerns morphological processing, Jackendoff & Audring (2016) follow the theory in which words are decomposed and retrieved at the same time, and also may strengthen one another (SCHREUDER & BAAYEN, 1995¹⁰ apud JACKENDOFF & AUDRING, 2016, p. 484; KUPERMAN, BERTRAM & BAAYEN, 2008¹¹ apud JACKENDOFF & AUDRING, 2016, p. 484)). The authors argue that “when the decompositional and whole-word strategies result in incompatible candidates, they compete. But when they result in compatible candidates, they reinforce each other, and their redundancy creates a more robust outcome and potentially a faster reaction time” (JACKENDOFF & AUDRING, 2016, p. 484).

They claim that schemas behave as words since they are lexical items in the relational morphology concerning the representational theory. As a result, they act as words in processing. The frequency of their use corresponds to their resting activation; accordingly, lexical items or schemas can prime them, which make their activation increase.

As a result, since Jackendoff & Audring (2016)’s Relational Morphology theory accounts for aspects of morphological analysis, processing and memory, it fits this research as a theoretical background since morphological analysis, processing and memory are the central issues discussed here.

2.2 Declarative/Procedural Model and the Second Language Acquisition

¹⁰ SCHREUDER, R. & BAAYEN, R. H. Modeling morphological processing. In: L. Feldman (Ed.), *Morphological aspects of language processing*. Hove: Erlbaum, p. 131–154, 1995.

¹¹ KUPERMAN, V.; BERTRAM, R. & BAAYEN, R. H. Morphological dynamics in compound processing *Language and Cognitive Processes*, v. 23, p. 1089–1132, 2008. doi: 10.1080/01690960802193688

Unlike Jackendoff & Audring (2016) that defend that lexicon and grammar does not need to be conceived separately, Ullman (2001a, 2001b) claims that is important to distinguish between mental lexicon and mental grammar when one is studying language. For him, while the lexicon is memorized, the grammar is a computational grammar. Therefore, the lexicon is part of the declarative system since it contains words – e.g. *table* – and idiomatic expressions – e.g. “*kick the bucket*” – that cannot be originated from operations of the grammar, and the grammar concerns the computational system, that is, the procedural system, as it derives meaningful linguistic forms – such as the past tense morpheme, *-ed*. This view is also contrary to the one defended by Jackendoff & Audring (2016) as they suggest that schema fits best the system instead of rule. Since schemas behave as words, they are part of the declarative representation instead of the procedural representation.

In addition, the Declarative/Procedural Model suggests that lexical memory is not constrained to information so that it can be retrieved to “multiple response systems”. As a result, after sequential and hierarchical structures are learned, they are stored in the left inferior parietal regions where it can be accessed by both the declarative and procedural systems.

Some researchers claim that the onset of language acquisition plays an important role in grammar acquisition (BIRDSONG, 1999¹² apud ULLMAN, 2001b, p. 108; JOHNSON & NEWPORT, 1989¹³ apud ULLMAN, 2001b, p. 108). As a result, Ullman (2001b) argues that a shift in the memory system happens in the L2 as it relies in the declarative rather than in the procedural memory for grammar acquisition:

the declarative/procedural model claims that in L1, the learning and use of grammar depend largely upon procedural memory, whereas the memorization and use of words depends upon declarative memory. In L2, by contrast, age-of-exposure sensitivity that affects grammatical computation is posited to involve the procedural memory system (p. 108).

Late learners of language may rely mostly on their declarative memory for grammatical computations, which differs from earlier learners, although some authors claim that declarative memory function starts to decay in the beginning of adulthood (BJORK & BJORK, 1996¹⁴ apud ULLMAN, 2001b, p. 109; KIRASIC, ALLEN, DOBSON, & BINDER,

¹² BIRDSONG, D. (ed.). *Second language acquisition and the critical period hypothesis*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 1999.

¹³ JOHNSON, J. S. & NEWPORT, E. L. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, v. 21 n. 1, p. 60-99., 1989.

¹⁴ BJORK, E. L. & BJORK, R. A. (eds.). *Memory*, v. 2. New York: Academic Press, 1996.

1996¹⁵ apud ULLMAN, 2001b, p. 109). As a result, linguistic forms may be stored in the lexicon, such as words, idioms, morphologically complex forms – e.g. *walked* – or even phrases and sentences. Lexical memory is capable of generalization in such a way that it can associate memorized patterns to new ones (PINKER, 1999¹⁶ apud ULLMAN, 2001b, p. 109; PRASADA & PINKER, 1993¹⁷ apud ULLMAN, 2001b, p. 109); hence, this lexicalization in the language may cause productivity, which does not mean that it is “unanalyzed and unstructured strings” (ULLMAN, 2001b). This construct seems to be in accordance with what Jackendoff & Audring (2016) defend in their Relational Morphology theory.

Even though some rules can be stored in declarative memory, they are not grammatical rules in the sense of being dependent on “procedural/grammatical computations”. Instead, they may be explicitly learned and explicitly accessed, but also implicitly used. (ULLMAN, 2001b).

Ullman (2001b) claims that age of exposure in L2 is a key factor in the “shift of dependence from procedural to declarative memory” (p. 110); nonetheless, this is not the only factor that leads to this dependence. Practice improves learning with procedural memory (SCHACTER & TULVING, 1994¹⁸ apud ULLMAN, 2001b, p. 110; SQUIRE *et al.*, 1993¹⁹ apud ULLMAN, 2001b, p. 110); subsequently, the declarative/procedural model argues that L2 practice can lead to an increase in the “dependence on procedural memory for grammatical computations” (ULLMAN, 2001b, p. 110). Ullman (2001b) claims that both practice and age of exposure affect not only dependence on procedural memory, but also grammatical proficiency. Despite the fact that Ullman (2001a, 2001b) has a different perspective from Jackendoff & Audring (2016) of how the mental lexicon and the mental grammar are stored in the mind, the former admits that the L2 bilinguals make use of their declarative knowledge to store grammatical rules.

2.3 Inflectional Morphemes studies

Most of times, “when L2 learners are engaged in message-oriented spontaneous speech or written production, they rely on their automatic competence rather than explicit knowledge” (JIANG, 2007, p. 7). In order to test automaticity in L2, it is important to reduce

¹⁵ KIRASIC, K. C.; ALLEN, G. L.; DOBSON, S. H. & BINDER, K. S. Aging, cognitive resources, and declarative learning. *Psychology and Aging*, v. 11, n. 4, p. 658-670, 1996.

¹⁶ PINKER, S. *Words and rules: the ingredients of language*. New York: Basic, 1999.

¹⁷ PRASADA, S. & PINKER, S. Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, v. 8, n. 1, p. 1-56, 1993.

¹⁸ SCHACTER, D. L. & TULVING, E. (eds.). *Memory systems 1994*. Cambridge, MA: MIT Press, 1994.

¹⁹ SQUIRE, L. R.; KNOWLTON, B. & MUSEN, G. The structure and organization of memory. *Annual Review of Psychology*, v. 44, p. 453-495, 1993.

the access to explicit knowledge since participants would make use of it instead of their implicit knowledge if the first one is available (BIRDSONG, 1989; SCHUTZE, 1996)²⁰. Experiments with the self-paced reading show that native speakers of English are extremely sensitive to ungrammatical sentences, since it can be observed a reading delay in sentences with grammatical violations (JIANG, 2007).

In what concerns on-line experiments with self-paced reading, both Jiang (2004b, 2007) and Carneiro (2011) have already carried out investigations regarding English inflectional morphemes. For that purpose, all sentences of the experiments conducted by Carneiro (2011) and Jiang (2004b, 2007) had a grammatical condition and an ungrammatical one; nevertheless, the two conditions never appeared in the same test form; therefore, their experiments had two different versions that were never exhibited to the same participant:

(11a) The new waitress sometimes tastes the dessert behind the back door.

(11b) *The new waitress sometimes taste the dessert behind the back door.

(CARNEIRO, 2011, p. 94)

Besides that, comprehension questions of the type yes-no followed some of the investigated sentences of Carneiro (2011) and Jiang (2004b, 2007) in order to verify if the participants were paying attention to what they were reading:

(12) After/ the patient/ recovered/ from the anesthesia/ the nurse/ call(ed)/ the surgeon/ in his mobile.

Did the patient recover? (CARNEIRO, 2011, p. 88)

(13) They enjoyed many of the debate(s) in their speech communication class.

Did the debates take place in the town hall meeting? (JIANG, 2007, p. 30)

Both studies report significant differences in between the processing cost of grammatical violations of native speakers and L2 learners of English. Notwithstanding, these studies found divergent results.

Jiang (2004b, 2007) examined English learners' competence through the self-paced reading in the moving-window paradigm in order to test the participants' sensitivity to grammatical violation. These experiments comprise the usage of regular plural (-s), verb subcategorization and verb *to be* agreement:

²⁰ The construct of implicit knowledge discussed above seems to be different from the one that Jackendoff & Audring (2016) defend; as the latter argue that morphological schemas rely on explicit knowledge rather than on implicit knowledge. However, we do not know if there is indeed strict separation between the two types of representations as N. Ellis (2005) defends that implicit and explicit knowledge interact.

(14) The tour guide tried to persuade all of the visitor(s) to taste the food. *Regular Plural (-s)* (JIANG, 2007, p. 30)

(15) The committee chair introduced/refused the speaker to everyone in the room. *Verb Subcategorization* (JIANG, 2007, p. 31)

(16) According to the speaker he/we is the first to have found this effect. *To be agreement* (JIANG, 2004b, p. 630)

The results suggest that natives were sensitive to all grammatical violations while L2 learners did not present substantial sensitivity regarding the plural morpheme (-s). These studies were conducted in the United States, and the data was gathered with Chinese English learners.

Jiang (2004b) argues that the nonnative learners' performance is related to the lack of integration of inflectional morphemes. As a result, there are two hypotheses involved here. Participants present either a deficit in their morphological representation or a mapping/processing problem in syntactic representations²¹

Carneiro (2011) compared the RTs of native speakers of English and Brazilian learners of ESL regarding the inflectional morphemes third person singular (-s) and past regular verbs (-ed) in an on-line experiment concerning the self-paced reading in the moving-window paradigm:

(17) The lazy schoolboy/ usually/ cause(s)/ the trouble/ in the Arts classroom. *Third Person Singular (-s) agreement* (CARNEIRO, 2011, p. 144)

(18) After/ the team/worked/ hard/ the coach/ earn(ed)/ the respect/ of his rivals. *Past Regular Verbs (-ed)* (CARNEIRO, 2011, p. 133)

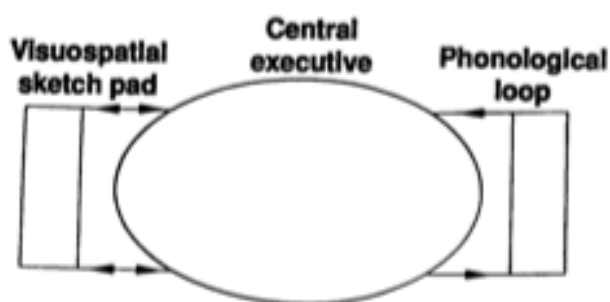
In this study, nonnative learners were not sensitive to English morphemes -ed and -s. Nonetheless, the native group examined by Carneiro (2011) did not show sensitivity to the absence of third person singular (-s) either. The latter may be due to the fact that all natives that took part in the experiments lived in Brazil and were familiar with the agreement violation made by Brazilians learners of English.

²¹ The last hypothesis is defended by Souza (2010) as it is plausible to assume that L2 learners have a specific interlingual mapping, which would explain the variability that L2 learners present in language production. This hypothesis is based on Jackendoff (1997), as it is in the lexicon that the representations are stored. As a result, it is in accordance with the Relational Morphology proposed by Jackendoff & Audring (2016) since schemas are part of the explicit knowledge of the language. However, Jiang (2004b) base his argument on the theory of Principle and Parameters proposed by Chomsky (1995).

2.4 Memory Studies

Baddeley (1992b) claims that working memory is related to brain processes that encompass temporary storage and manipulation of information so that it is possible to conceive cognitive tasks such as “language comprehension, learning, and reasoning” (p. 556). As a matter of fact, the author separates working memory into three subcomponents: central-executive, visuospatial sketchpad, and phonological loop. The first one is responsible for attention-controlling and the other two are subordinate to it. Visuospatial sketchpad controls visual images while phonological loop is essential for acquisition of the native language and also L2 vocabulary because it deals with storage and information rehearsal.

Figure 1: Baddeley’s (1992b) working memory model p. 557



Furthermore, Baddeley (2015) states that despite the fact that there are no settled methods for measuring working memory capacity, he tends to follow Cowan’s (2005, 2010) proposal in which the WM capacity has a “maximum of five episodes or chunks” (BADDELEY, 2015, p. 26).

However, as pointed out by Wood *et al.* (2001) this model may have many problems concerning its division, as it is very hard to find and divide the central executive, which results in the fragmentation of it. Consequently, they believe that the model proposed by Salthouse & Babcock (1991) is the one which best suits the memory battery test that they conducted. Salthouse & Babcock (1991) carried out experiments with a wide range of adults – from 18 to 87 years old – to verify if working memory is affected by aging. The authors have discovered that processing speed is the component that suffers the biggest impact with aging. Additionally, processing efficiency is closely related to speed.

Salthouse & Babcock (1991) defend that it is important to see “working memory not as a single discrete structure, but rather as a dynamic interchange among three conceptually

distinct aspects or components—processing efficiency, storage capacity, and coordination effectiveness” (p. 763). As a consequence, working memory concerns the construct of retaining information as the same or other information is being processed (BADDELEY, 1986²² apud SALTHOUSE & BABCOCK, 1991, p. 763; CARPENTER & JUST, 1989²³ apud SALTHOUSE & BABCOCK, 1991, p. 763; SALTHOUSE, 1990²⁴ apud SALTHOUSE & BABCOCK, 1991, p. 763). As a consequence, McDonald’s (2006, 2008b) tasks may be used to exemplify this since their grammaticality judgments with memory load require participants to store numbers while classifying the grammaticality of sentences that were orally presented.

Wood *et al.* (2001) adapted a working memory battery test (BAMT-UFMG) considering the model proposed by Salthouse & Babcock (1991). Their BAMT-UFMG was done in pencil and paper, and it consists of a simple task in which any literate person can perform. BAMT-UFMG involves two sets of similar tasks that cover numerical and verbal stimuli that measure every component from working memory – coordination capacity, storage capacity and processing efficiency. Despite the fact that the tasks were simple, the processing load was progressively increased up to the limit of processing capacity. The data was collected with students from elementary and high school, Psychology, CPOR (*Centro de Preparação de Oficiais da Reserva*), and elderly people that had been trained to do similar tasks as the BAMT-UFMG and others who had not.

Other studies about memory have been carried out in Brazil. While Finardi (2009) and Azevedo (2012) investigated how working memory affects L2 learners’ minds, Jesus (2012) examined the role of declarative and procedural memory in bilingual systems. Both Finardi (2009) and Azevedo (2012) analyzed if WMC affected the phonological performance of L2 learners, and Jesus (2012) if declarative and procedural memory and proficiency interacted. Finardi (2009) found results proving that WMC correlates with L2 speech acquisition of complex syntactic structure. Azevedo’s (2012) research discovered evidence that supports the premise that WMC is related to listening comprehension performance. Besides, Jesus’s (2012) results indicate that high proficient bilinguals are more accurate than low proficient bilinguals and monolinguals in tasks that measure declarative and procedural memory. Although Azevedo’s (2012) participants were low proficient bilinguals, she found results that

²² BADDELEY, A. D. *Working memory*. Oxford, England: Clarendon Press. 1986.

²³ CARPENTER, P. A., & JUST, M. A. The role of working memory in language comprehension. In: D. KLAHR & K. KOTOVSKY (Eds.), *Complex information processing: The impact of Herbert A. Simon* Hillsdale, NJ: Erlbaum. p. 31-68, 1989.

²⁴ SALTHOUSE, T. A. Working memory as a processing resource in cognitive aging. *Developmental Review*, v. 10, p. 101-124, 1990.

associate bilingualism and WMC. This may be due to the fact that bilingualism may offer a cognitive advantage (BIALYSTOK *et al.*, 2009).

CHAPTER 3

METHODOLOGY

3.1 Overview

This study was conducted over two experimental sessions. One of them was performed by participants in person and the other one online. The face-to-face session was carried out first. In that session, participants performed an off-line acceptability judgment (AJ) task in two formats, one with a time constraint and another one with memory load. Afterwards, a n-back task, which demanded that participants selected similar items showed to them in a sequence of letters, was performed. Once participants have completed this section, they answered a web-based version of the Vocabulary Levels Test (VLT), adapted from Nation (1990).

The AJ tasks verified L2 learners of English acceptability to sentence violations regarding regular past (-*ed*) and third person singular (-*s*) morphemes; therefore, it was examined if participants were sensitive to the absence of these morphemes in contexts that they were expected to be. Moreover, this task had two formats: a speeded acceptability judgment similar to the one investigated by Souza & Silva (2015) and Silva (2016), and also one with an external memory load such as the one McDonald (2006, 2008b) carried out. As a consequence, every participant took the two formats of the AJ. Correspondingly, the format with a time constraint demanded that participants ranked the sentences in up to 6000 milliseconds (ms), and the one with memory load asked that participants memorized a random seven-digit number while they judged the sentences.

N-back was carried out by participants in the 2-back load, correspondingly to the one conducted by Kane *et al.* (2007). The task had four target lists of 48 letters, in which each block contained eight targets and 40 foils. Following Kane *et al.* (2007), we instructed the participants to judge as fast and as accurately as possible whether a letter matched the one presented 2 times ago.

It is important to highlight that all experiments were performed in the same session and always in the same order, so that there would not be any effect considering when and the order that the participants carried out the experiments. They were conducted at the Psycholinguistics Lab at *Universidade Federal de Minas* on a portable laptop using the

software PsychoPy²⁵. According to Pierce (2009), the software PsychoPy is an open source and a trustworthy tool in carrying out experiments with visual and auditory stimuli, as it also enables a wide range of design possibilities. Therefore, it was used in this research to run the experiments. This software enables to randomize the stimuli; thus, each participant saw the sentences in the AJ tasks in a different order so that there would be no effect considering the position of the sentence in the experiment.

The VLT was used to determine the level of proficiency of participants, according to the investigation conducted by Souza & Silva (2015) and Silva (2016). Nonetheless, the task was adapted to be done in an online platform instead of on paper²⁶. Based on the VLT scores, L2 participants were divided into two groups, low and high proficient, and their accomplishments were compared in the AJ tasks, in all the two formats, to verify if the two groups presented different results concerning the grammatical and ungrammatical sentences:

Table 1. Structure of the research

Experiment 1 – Acceptability Judgment (AJT) Task with time restriction	in-person
Experiment 2 – Acceptability Judgment Task (AJWM) with external memory load	
Experiment 3 – N-back task	
Vocabulary Levels Test (VLT)	online

Afterwards, all the data was statistically treated in order to check if there was any significant difference between the different groups in each task, such as for instance, the high and low proficient L2 learners, and so forth. Thus, the RT and performance of every group was statistically verified in order to delve into if the results of every different group were the outcome of factors such as proficiency, experimental paradigm, memory load and processing capacity.

3.2 Experiment 1: Acceptability Judgment with Time Restriction (AJT)

The objective of the first experiment was to investigate if a time constraint would pose a processing challenge to the two groups investigated, high and low proficient, considering the outcome of grammatical and ungrammatical sentences' RT and score. In order to do so,

²⁵ Available at < <http://www.psychopy.org>>.

²⁶ Participants were given an access code to do VLT. The test is available online at <<http://www.classroomclipboard.com/635396/Home/Test/72084A4F63844B658F9BEE0611A2859F>>. See Appendix 1.

participants had up to 6000 msc, following Souza & Silva (2015) and Silva (2016), to read every sentence as a whole. Each sentence – targets and distractors – had 40 characters at most including spaces. As a consequence, participant’s RT and sentence score were recorded when classifying grammatical and ungrammatical sentences (Appendix 2). The participants had to rank the sentences from 1 to 5 following the Likert scale.

From all the forty experimental sentences, ten of them were the targets and the other ones were distractors. There were grammatical and ungrammatical sentences concerning the morpheme –ed – past regular verbs –, and the morpheme –s – third person singular agreement²⁷:

(19a) Man land on the moon in 1969. *Absence of morpheme -ed*

(19b) Robert phoned the police yesterday. *Morpheme -ed*

(20a) He love getting up at 10 o’clock. *Absence of morpheme -s*

(20b) My uncle eats breakfast every morning. *Morpheme -s*

In order to get familiar with the experiment, participants conducted a training of six sentences following the same parameters adopted in the trial (Appendix 2).

3.3 Experiment 2: Acceptability Judgment with Memory Load (AJWM)

The second acceptability judgment’s aim was to examine if individuals with higher working memory capacity had a higher ability to perceive the morpheme violations: –ed – past regular verbs – and –s – third person singular agreement –, when their working memory capacity was exhausted. This task had no time restriction and demanded that participants memorized a random seven-digit number²⁸ before classifying each sentence. After classifying the sentences, the participant had to digit the numbers (Appendix 3).

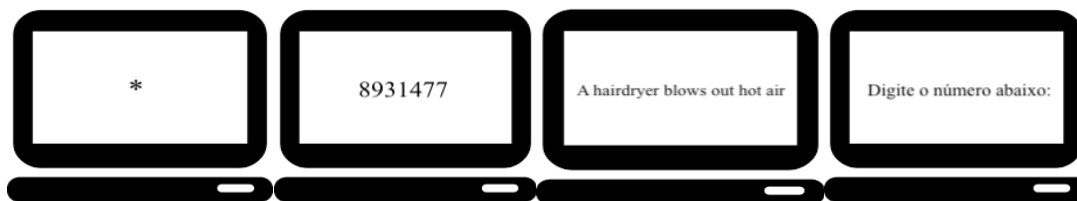
Although they could take as much time as needed, they were instructed to do as fast as they could. An asterisk was presented for 500 msc before each number to orient the participants, the number for 1500 msc and after it the sentence, following the work done by McDonald (2008b). Unlike the research conducted by McDonald (2008b), all the stimuli – numbers and sentences – were visually presented to the participants in a computer screen. Furthermore, participants were instructed to type the numbers after they classified the

²⁷ All the sentences used in the experiment were either taken or adapted from Carneiro (2008) and Carneiro (2011).

²⁸ It is important to state that the numbers were randomized in a website, <<https://www.random.org/strings/>>, so that they could be randomly distributed.

sentence; thus, all answers were registered by the software PsychoPy:

(21) Representation of the AJWM experiment²⁹



Such as in the acceptability judgment with time restriction, this experiment had forty sentences from which ten were the targets concerning the –ed morpheme and the –s morpheme³⁰. Besides, the sentences – targets and distractors – had up to 52 characters including spaces. Despite the fact that McDonald (2008b) adopted a classification parameter of 0, meaning incorrect, and 1, correct, the Likert scale from 1 to 5, adopted by Souza & Silva (2015) and Silva (2016), as described in the previous experiment, was implemented. Before performing the trial sentences, participants conducted a training of ten sentences (Appendix 3).

3.4 Experiment 3: 2-back

The 2-back experiment's purpose was to compare high proficient and low proficient participants' RT and also check their performance while performing an experiment that demanded working memory capacity. As a consequence, we aimed to check if proficiency influences WMC.

Participants had to choose between different and corresponding letters in four separated lists, such as in Kane *et al.* (2007) (Appendix 4)³¹. In the trial, the letters were arranged in a way that upper and lower cases were placed after one another, despite the fact that in Kane *et al.*'s. (2007) study the letters format was arbitrary³². Before the trial, participants conducted three lists of training (Appendix 4). Participants were instructed to carry out the experiment as fast as possible. An orienting asterisk was shown for 500 msec

²⁹ It is important to state that all instructions were presented to participants in Portuguese; thus, this is the reason why the last screen appears in Portuguese.

³⁰ Such as in the previous experiment, sentences were taken or adapted from Carneiro (2008) and Carneiro (2011).

³¹ This experiment implemented the same lists, in both training and trial, developed by Kane *et al.* (2007) for their 2-back experiment. Despite the fact that Kane *et al.*'s. (2007) 2-back task had eight lists, this research used only seven of them.

³² We decided to follow this format because our participants experienced confusion in the training when two letters, such as in lower case, were exhibited in sequence. Besides, participants had already carried out other experiments when they performed this experiment; therefore, we intended to avoid confusion.

before each list and a break of 3000 msc was placed between every list. Participants were oriented to press 1 every time a letter repeated after two others and 3 when there was no correspondence:

(22) Representation of the 2-back experiment



3.5.1 Materials

All the studies upon which this study is based are described in this section. Consequently, it is divided by the tasks that participants were supposed to carry out. First, the experiments performed at the lab are explained, and afterwards, the task conducted online. Furthermore, there is a comprehensive report about the findings that were obtained by different authors using these tasks, as well as the advantages and disadvantages of using such tasks.

3.5.2 Acceptability judgment (AJ) task

Some authors adopt grammaticality judgment tasks – such as McDonald (2006, 2008b), Souza; Soares-Silva & Silva (2016), Souza & Oliveira (2017), among others – while others adopt acceptability judgment tasks – such Souza *et. al.* (2015), Souza & Silva (2015); Silva (2016), among others – in their studies. The difference between acceptability and grammaticality judgment is a controversial issue discussed in Bard, Robertson & Sorace (1996), Souza *et. al.* (2015), and Silva (2016). Bard, Robertson & Sorace (1996) argues that while the grammaticality concerns the theoretical linguistic framework, the acceptability is about the learner’s perception about the input received. Therefore, they defend that the “acceptability judgment [...] is the speaker’s response to the linguist’s inquiries” (p. 33).

Taking Chomsky’s (1957) well-known sentence “Colorless green ideas sleep furiously” as an instance, it is possible to realize that despite the fact that this sentence is grammatically acceptable, it does not make any sense. Chomsky (1957) states that “the notion “grammatical” cannot be identified with “meaningful” or “significant” in any semantic sense” (p. 15). As a result, this is the reason why we decided to adopt the acceptability judgment instead of grammaticality judgment.

3.5.2.1 Acceptability judgment task with time restriction (AJT)

Grammaticality judgment tasks are considered by many authors an off-line measure since they register explicit and declarative knowledge (SOUZA & OLIVEIRA, 2017). In what concerns tests that focus on grammatical acceptability, the access to explicit knowledge could occur especially in tasks where the time is not limited (SOUZA & SILVA, 2015; SOUZA & OLIVEIRA, 2017):

higher amounts of time for judgment calls would serve as a task tapping into explicit, declarative knowledge about language facts and more controlled processing, both of which would be unavailable should the task be administered in a mode that gave participants but few seconds for their responses (usually 4 to 6 seconds). In such strictly timed grammaticality judgment tasks, the measure would be of implicit linguistic knowledge (SOUZA & OLIVEIRA, 2017, p. 1688).

High proficient L2 learners exhibit a higher capacity to tolerate some constructions that monolinguals would face difficulty to process or accept. Fernández & Souza (2016), Souza (2014) and Fernández; Souza & Carando (2017) have attested this phenomenon through tasks in which language processes in comprehension and production were demanded from participants. The authors claim that L2 learners are able to innovate their linguistic representations in both L1 and L2.

Nonetheless, Souza; Soares-Silva & Silva (2016) could not find the same level of tolerance to L1 argument structure constructions similar to L2 ones in a timed grammaticality judgment task. The participants of their study were Brazilians who were highly proficient English bilinguals. The researchers propose that bilingual cross-linguistic effects are too weak to be detected in a speeded judgment task. Furthermore, other researchers argue that acceptability/grammaticality judgment tasks, irrespective of being timed or not, can only access explicit knowledge (VAFEE *et al.*, 2016³³ apud SOUZA & OLIVEIRA, 2017, p. 1689; KIM & NAM, 2016³⁴ apud SOUZA & OLIVEIRA, 2017, p. 1689).

As a consequence, Souza & Oliveira (2017) investigated if a grammaticality task with time constraint and a maze task, an online processing task, were measuring different constructs. The researchers gathered data from high proficient Brazilian learners of English and Brazilian monolingual speakers of Portuguese. Even though bilinguals and monolinguals behaved differently in the maze task, bilinguals and monolinguals had a similar behavior in

³³ VAFEE; P.; SUZUKI, Y. & KACHISNKE, I. Validating grammaticality judgment tests – Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, Cambridge University Press, p. 1-37, 2016.

³⁴ KIM, J. & NAM. H. Measures of implicit knowledge revisited – Processing modes, time pressure, and modalities. *Studies in Second Language Acquisition*, Cambridge University Press, p. 1-27, 2016. <https://doi.org/10.1017/S0272263115000510>

the grammaticality judgment. This suggests that there was no influence of the L2 over the L1 when judging the experimental sentences.

In order to carry out the first experiment, we followed Souza & Silva (2015) and Silva (2016), who conducted a speeded acceptability judgment, so that we could verify the participants' acceptability in relation to grammatical and ungrammatical sentences. Many researchers adopt the Likert scale in AJ tasks since it is a scale of measure that is easily implemented and provide useful information (WESKOTT & FANSELOW, 2011³⁵ apud OLIVEIRA, 2016, p. 105; FUKUDA *et al.*, 2012³⁶ apud OLIVEIRA, 2016, p. 105, SOUZA & OLIVEIRA, 2014). As a consequence, participants were asked to rank the sentences from 1 to 5, following the Likert scale, according to its acceptability:

Table 2. Acceptability Judgment Task Likert Scale (SILVA, 2016, p. 61)

Numeric keypad	Judgment levels
1	totally unacceptable
2	not well-formed, almost unacceptable
3	not well-formed, but maybe acceptable
4	slightly ill-formed, almost perfect
5	totally perfect

The types of sentence violations involved in the experiments investigated by Souza & Silva (2015) and Silva (2016) are: transitivized unergative verb, subject-verb agreement, and Wh-movement:

(23) *The man laughed the children during the party. *Transitivized unergative verb* (SOUZA & SILVA, 2015 p. 197).

(24) *The girl give the cats milk twice a day. *Agreement violation* (SOUZA & SILVA, 2015 p. 197).

³⁵ WESKOTT, T.; FANSELOW, G. On the informativity of different measures of linguistic acceptability. *Language*, v. 87, p. 249-73, 2011.

³⁶ FUKUDA, S.; GOODALL, Grant; MICHEL, Dan; BEECHER, Henry. *Is magnitude estimation worth the trouble?* In: CHOI, J.; HOGUE, A.; PUNSKE, J.; TAT, D.; SCHERTZ, J.; TRUMAN, A. Proceedings of the 29th West Coast Conference on Formal Linguistics, Somerville: Cascadilla Proceedings Project, p. 328- 336, 2012.

(25) *What did Steven read the book that Helen talked about? *WH movement violation* (SOUZA & SILVA, 2015 p. 197).

Based on the means of low proficient and high proficient test takers, Souza & Silva (2015) and Silva (2016) estimated that Brazilian learners of English would take up to 6000 msc to classify a sentence written in English with about 40 characters, not including the spaces. Consequently, we followed the same patterns adopted by Souza & Silva (2015) and Silva (2016) in the speeded acceptability judgment task.

3.5.2.2 Acceptability judgment task with external memory load

McDonald (2006) carried out a study to examine if late English bilingual learners' grammatical knowledge deficit could be explained by: (i) poor memory capacity, (ii) low L2 decoding and/or (iii) slow L2 processing speed. McDonald (2006) conducted two experiments, the first one compared the performance of late bilinguals to unstressed L1 learners, and the second one investigated the performance of native speakers who had to perform tasks under stress.

In the first experiment, participants were asked to carry out four tasks: (i) a size judgment span, which evaluated working memory span, (ii) a gating task, which investigated decoding ability, (iii) a word detection process, which checked speed of processing, and (iv) a grammaticality judgment task, orally presented. Besides performing the tasks described above, L2 participants had to perform a task that measured grammatical performance.

The grammaticality judgment task investigated: (i) definite article, (ii) subject–verb–object word order, (iii) wh-questions, (iv) yes/no questions, (v) present progressive, (vi) third person singular agreement, (vii) plurality agreement, and (viii) past tense agreement. Both plurality and past tense were investigated in its regular and irregular forms:

(26) The lady drove the/– same car for the past 20 years. *Definite Article* (MCDONALD, 2006, p. 399)

(27) The young child likes carrots/likes carrots the young child. *Word order* (MCDONALD, 2006, p. 399)

(28) What do/– you think about the new coach? *Wh- questions* (MCDONALD, 2006, p. 399)

(29) Does the teacher drive/drives the teacher a really fancy red car? *Yes-No question*

(MCDONALD, 2006, p. 399)

(30) The little girl is playing/play with her dolls. *Present Progressive* (MCDONALD, 2006, p. 399)

(31) The boy jumps/jump whenever he is startled. *Third person singular* (MCDONALD, 2006, p. 399)

(32) This meal takes more than four pans/pan to prepare. *Regular Plural* (MCDONALD, 2006, p. 399)

(33) After dancing for 2 hours, my feet/foots began to hurt. *Irregular Plural* (MCDONALD, 2006, p. 399)

(34) Last night my friend walked/walk home after dark. *Regular Past* (MCDONALD, 2006, p. 399)

(35) Yesterday morning he took/taked the car for a test drive. *Irregular Past* (MCDONALD, 2006, p. 399)

The author instructed participants to press 1, when the sentence had no grammatical mistake, and 0, when it was ungrammatical. Furthermore, a language background questionnaire was administered to participants. Participants comprised a group of 50 native speakers and 50 late bilinguals. The late L2 learners moved to the U.S.A. when they were 12 year-old or after it.

In the second experiment three hundred English native speakers had to perform the same tasks from the first experiment, but with different types of stressor, such as low memory stress – memorize a four digit number –, high memory stress – memorize a seven digit number –, noise stress, response deadline stress, and compressed speech. The grammaticality judgment task was the task that demanded that participants memorized numbers.

The L2 learners exhibited poor ability in relation to the following: working memory span, gating scores, word detection latencies, and grammaticality judgment accuracy. However, McDonald (2006) argues that native speakers who face problems in memory capacity, language decoding, and speed processing also present problems in the tasks conducted in this study.

Furthermore, McDonald (2008b) investigated grammaticality processing of (i) subject-verb-object (SVO) order, (ii) regular plural morpheme (-s), and (iii) third-person

subject-verb agreement morpheme (-s) in two off-line experiments in the format of grammaticality judgment tasks. Even though both of them had no temporal restriction, one had memory load:

(36) *Opens the girl the door. *Subject-verb-object (SVO) order* (MCDONALD, 2008b, p. 982)

(37) The shopper needed a cart(s). *Regular plural morpheme (-s)* (MCDONALD, 2008b, p. 982)

(38) A dog chase(s) two balls. *Third-person subject-verb agreement morpheme (-s)* (MCDONALD, 2008b, p. 980)

Participants were adult English native speakers, and half of them performed the experiment with external memory load, and the other half without it. Hence, it was demanded from the group with memory load that they memorized a random seven-digit sequence before judging the sentences. An asterisk was used to orient participants for 500 msc, which was followed by the number that was presented for 1500 msc. All of McDonald's (2008b) sentences were introduced through auditory stimuli, and participants were asked to classify the sentences according to its grammatical accuracy; there were only two options, they would be categorized as either 1, correct, or 0, incorrect.

Such as Jiang (2004b, 2007) and Carneiro (2011), McDonald (2008b) had different lists with different conditions concerning grammaticality – a grammatical condition and an ungrammatical condition; however, the different conditions of the same sentence were never presented to the same participant:

(39a) The girl pads the pillow

(39b) The girl pad the pillow (MCDONALD, 2008b, p. 982)

Additionally, a memory capacity test was previously administered to all participants. After the grammaticality judgment test was conducted, participants did a phonological ability task. According to McDonald (2008b), “all effects were localized to ungrammatical, rather than grammatical, sentences. This is probably because the grammaticality judgment task is viewed as an error-detection task, in which the default response is to say “correct” unless an overt error is detected” (p. 982). Accordingly, it was attested that participants with higher memory capacity and higher phonological ability could better detect mistakes. External memory load showed no effect for SVO order, and only marginally effect for plural (-s);

however, it had substantial effect for subject-verb agreement (-s). Therefore, McDonald (2008b) proposes that memory capacity affects subject-verb agreement processing. Consequently, low memory span and memory load influenced the capacity to process subject-verb agreement.

As a result, this study adapted the models proposed by McDonald (2006, 2008b), Souza & Silva (2015) and Silva (2016) in what concerns the judgment of acceptability, so that both formats can be investigated, the speeded format and the one with external memory load.

3.5.3 N-back

Kane, Conway, Miura & Colflesh (2007) have conducted an experiment in which they evaluated N-back's validity in comparison to a working memory span and a Ravens Advanced Progressive Matrices Test (RAPM).

N-back is a type of task that demands that participants select every stimulus in a sequence that corresponds to the one that was shown n times previously. For that purpose, they conducted tasks in which 2-back ($a-b-c-b$) and 3-back ($a-b-c-a$) were presented to the participants. According to the authors, N-back has not been very much empirically validated. Nonetheless, in their research “ n -back has face validity as a WM task because participants must maintain and update a dynamic rehearsal set while responding to each item” (p. 615).

Working memory capacity encompasses the construct of executive control over interference or competition (ENGLE & KANE, 2004; MAY *et al.*, 1999); thus, Kane *et al.* (2007) examined if n-back sequences could be used to investigate if interference affects performance. Nonetheless, Roberts & Gibson (2002) claim that n-back tasks comprise features of simple short-term memory span rather than complex WM span. Other researchers such as Oberauer (2005) and Gray *et al.* (2003) examined n-back's relation to complex WM tasks and RAPM correspondingly. While the first found weak-to-moderate relations, the second found strong relations.

As Meule (2017) pointed out, it is important to take into consideration whether each study is investigating n-back regarding accuracy, reaction time or number of errors. Most of times, RT does not correlate with accuracy (CARTER *et al.*, 1998³⁷ apud MEULE, 2017, p. 1). Accordingly, a greater number of errors are related to greater reaction times. Thus, Meule (2017) claims that “reaction times and accuracy in n -back tasks should not be interpreted

³⁷ CARTER C. S.; PERLSTEIN, W.; GANGULI, R.; BRAR, J.; MINTUN, M., & COHEN, J. D. Functional hypofrontality and working memory dysfunction in schizophrenia. *Am. J. Psychiatry*, v. 155, p. 1285–1287, 1998. doi: 10.1176/ajp.155.9.1285

interchangeably” (p. 2). He claims that different experiments ought not to reach analogous conclusions. The investigations conducted by Kopf *et al.* (2013), which studied emotional words, and by Marx *et al.* (2011), which studied pictures of emotional scenes, discovered that accuracy was a key factor for their emotional stimuli while Cromheeke & Mueller (2016), who studied faces, discovered that RT had essential effects for their emotional stimuli. Even though these studies had different measures of interpretation, the latter concluded that working memory was enhanced due to focus on affective information.

For their research, Kane, Conway, Miura & Colflesh (2007) used accuracy as their measure. Hence, they selected eight phonologically different letters for the experiment, *B, F, K, H, M, Q, R, X*. The experiments were conducted with memory load of 2-back and 3-back. For every n-back condition there were four lists of 48 letters. The chosen letters appeared 6 times in a list when they were the targets. Every block contained eight targets, which represented 16.67% of the trials, and 40 foils, which represented 83.33% of the trials. Beforehand, every participant had to complete a 2-back and 3-back practice with 40 trials each. Afterwards, they performed a practice that alternated between 2- and 3-back with eight blocks containing 48 trials each. They designed the trials in a way that letters were shown arbitrarily in upper and lower case in order to avoid identification due to just perceptual features. Participants were instructed to press either 1, when the letters matched the condition, and 3, when they did not.

As a result, this study employed the same lists developed by Kane *et al.* (2007), in what concerns the 2-back condition, so that we could have a working memory capacity metric, and participants’ WMC could be evaluated. Despite the fact that Kane *et al.* (2007) only used accuracy as their measure, this study also investigated participants’ RT so that it could be checked if there was any statistical difference when the results of high and low proficient participants were compared in relation to their performance and RT.

3.5.4 Vocabulary Levels Test (VLT)

“Proficiency in a second language (L2) may be generally understood as the ability of fluent use of that language” (SOUZA & SILVA, 2015 p. 189). However, this definition may be much more complicated than that since there are many factors involved in this construct, in other words, proficiency is, among many things, related to one’s capacity to properly and accurate use a language. According to Alderson (2005), language ability is largely related to vocabulary size; thus, one’s performance in a language can be measured considering the extent of the vocabulary.

Moreover, in what concerns vocabulary knowledge, a division between “vocabulary size” and “vocabulary depth” is frequently made. The former involves the amount of words an individual knows, as well as the form and the meaning of words, and the latter concerns the extent to which somebody is acquainted with the word, e.g. “collocation, derivative forms, polysemous meaning senses and etc.” (SCHIMITT, 2014, p. 916).

Souza & Silva (2015) and Silva (2016) compared the Vocabulary Levels Test (VLT) to the Oxford Placement Test (OPT) in order to attest the validity of the first one. The tests were used to measure the performance of Brazilian learners of English as an L2 that were in college so that VLT could be a reliable tool for proficiency assessment concerning college level students in Brazil.

VLT has five distinct parts in which the participants have to match three definitions to three out of six options:

Figure 2: Matching in VLT (SOUZA & SILVA, 2015, p. 193)

1 – business	
2 – clock	() part of a house
3 – horse	() animal with four legs
4 – pencil	() something used for writing
5 – shock	
6 – wall	

In addition, participants needed to score 12 out of 18 items to proceed to the following section (NATION, 1990). Although Silva (2016) decided to implement a minimum of 14 items per section, Souza; Duarte & Berg (2015) did not find any substantial difference between 12, 13 and 14 items; therefore, this research decided to implement the 12 points threshold.

Besides, the vocabulary size examined on VLT corresponds to lemmas, that is, word families. The test takers are divided according to their proficiency in five levels that match CEFR (Common European Framework of Reference for Languages) (SOUZA & SILVA, 2015):

the VLT estimates vocabulary size levels by correspondence between level and word frequencies bands based on the Brown Corpus. Successful completion of level 1 corresponds to knowledge of the 2,000 most frequent words; completion of level 2 corresponds to the 3,000 most frequent words, level 3 corresponds to the 5,000 most frequent words, level 4 is a special section corresponding to academic and scientific vocabulary, and level 5 corresponds to knowledge of the 10,000 most frequent words (SOUZA & SILVA, 2015, p. 193).

Souza; Duarte & Berg (2015) could not find any defining effect for level 4. The authors interpreted this finding as a result of the fact that this section was constituted of academic and scientific vocabulary, which is largely formed by words originated from Latin; thus, they are cognates with Portuguese words. Since level 4 was discarded, participants from the study of Souza & Silva (2015) had to reach level 5 in order to be considered high proficient as well as attain 91% in OPT.

There was no time constraint to perform VLT in Nation's (1990) original design. Nonetheless, a reasonable and significant correlation between RT and vocabulary size was found by Laufer & Nation (2001). Consequently, Souza; Duarte & Berg (2015) demanded that participants completed the test in up to 10 minutes. Silva (2016) claims that this time restriction fits the test since level 4 was not used to determine proficiency by Souza; Duarte & Berg (2015). Besides that, Silva (2016) argues that "by introducing the speed limit, the non-discriminatory section may be functioning as a modulator of how far the test-taker will be likely to reach in face of the temporal constraint for execution of the test's task" (p. 52).

Since Souza & Silva (2015) managed to validate the VLT in relation to the OPT, VLT was used in this research to obtain L2 proficiency level because it is a trustworthy tool to determine the proficiency of Brazilian learners of English as an L2.

CHAPTER 4

DATA ANALYSIS

4.1 Statistical analysis

For each experimental analysis, descriptive statistics will be presented before inferential statistics concerning the tests carried out.

Before the analysis of each experiment, the Shapiro-Wilk test was conducted so that the normality of participants', both high and low proficient, RT and performance could be verified. As a consequence, when the results were not normally distributed, we used the non-parametric test Mann-Whitney, and when they were normally distributed, we conducted the analysis with the unpaired t test, which is a parametric test.

The AJ tasks were statistically analyzed considering the grammatical and ungrammatical sentences of high and low proficient bilinguals separately. Therefore, high and low proficient bilinguals were divided into two groups and their results in both grammatical and ungrammatical sentences were compared. In the 2-back task, the performance of high and low proficient participants was contrasted.

In the AJ task with memory load, participants were supposed to memorize a seven-digit number as best as they could. In order to discover how many digits they could retain and how different the sequences that they typed were from the correct ones, we used the Levenshtein Distance, which is a distance metric that enables to quantify the variation between two sequences. Therefore, it is possible to account for all the modifications, such as omissions, insertions or substitutions, made by the participants (LEVENSHTEIN, 1966).

Furthermore, a ROC curve (Receiver Operating Characteristic curve) was also conducted in the memory load of the AJ task and in the 2-back, so that a diagnosis could be drawn in the performance of both high and low proficient participants considering their working memory capacity. We aimed to check if the WMC was correlated to the level of proficiency.

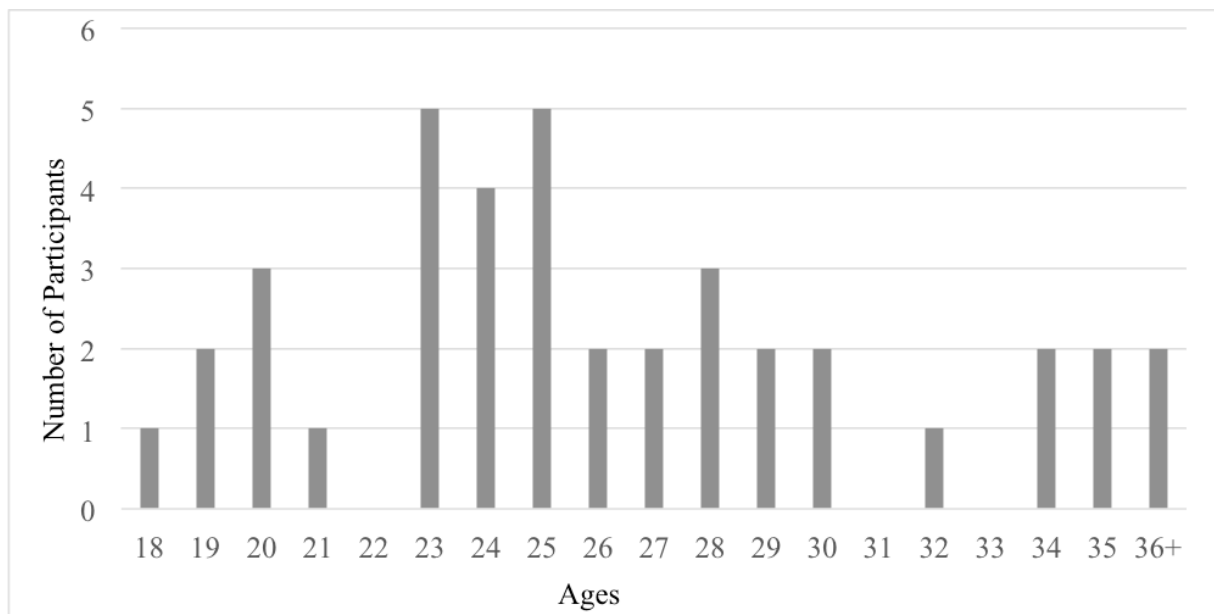
4.2 Participants

All participants were selected based on the fact that their native language was Brazilian Portuguese and English was their second most dominant language. Therefore, if one had learned English as an L2, but another language, such as French or German for instance, became their second most dominant language, the person could not take part in the

experiments.

Furthermore, participants could not be more than forty-five years old. Salthouse (2009) claims that “cognitive decline begins shortly after individuals reach maturity” (p. 508). However, Albert & Heaton (1988) and Ronnlund *et al.* (2005) argue that performance starts to decline when people reach 50 and 55 years old respectively. Additionally, Salthouse (2009) found evidence proving that the peak of cognitive performance is between the ages 22 and 27, followed by the ages 27 to 42. Based on that, the age range from the participants varied from 18 to 43:

Figure 3: Participants’ ages



Formal education is also a factor accounted for many researchers such as Salthouse (2009); therefore, participants needed to be, at least, pursuing a college degree to qualify as a candidate for the present research.

Participants were divided into two groups, high and low proficient, concerning their level of proficiency obtained through VLT. In order to be considered high proficient, participants needed to score level 5 in VLT, and the other levels were considered low proficient. The high proficient group comprised twenty-four participants while the low proficient group had fifteen participants. The data was gathered from forty people, nonetheless, the results from one low proficient participant had to be eliminated, as he was not able to score the grammatical target sentences in the AJT task. It is important to highlight that all participants carried out all three experiments always in the same order.

4.3 Experiment 1: Acceptability Judgment with Time Restriction

First, we present the results of the two groups', high and low proficient, RT and then their sentence score in grammatical and ungrammatical sentences – considering the morphemes –ed, regular past, and –s, third person agreement under investigation³⁸. Below we find the means and SD of the grammatical and ungrammatical sentences of the high proficient and low proficient group's RT:

Table 3: Means and SD of Participants' RT in the AJT task

Sentence type	Means	SD
Grammatical high proficient	2,490	570
Ungrammatical high proficient	2,690	570
Grammatical low proficient	3,050	520
Ungrammatical low proficient	3,230	480

Afterwards, the normality test Shapiro-Wilk was carried out in order to check if participants' RT and score in the target sentences were normally distributed. The value .05 was fixed to reject the null hypothesis.

When we compared the RT of high proficient participants in both grammatical and ungrammatical sentences, the results showed that they were not normally distributed; however, the low proficient participants' RT was normally distributed:

Table 4: Shapiro-Wilk test for Participants' RT distribution in AJT

Group	W	p value
High proficient group: grammatical and ungrammatical sentences	.953	5.70E-07
Low proficient group: grammatical and ungrammatical sentences	.987	.171

Since not all the data was normally distributed, we decided to analyze it with the non-parametric test Mann-Whitney. The results indicate that when the high proficient group's RT in grammatical and ungrammatical sentences were compared, it was not significant different ($U= 242$, $Z= -0.93819$, $p> .34722$). The same is observed in the RT of the low proficient group, in grammatical and ungrammatical sentences ($U= 87$, $Z= -1.03695$, $p> .29834$).

Consequently, we did the same analysis for the sentence score of the AJT task. The

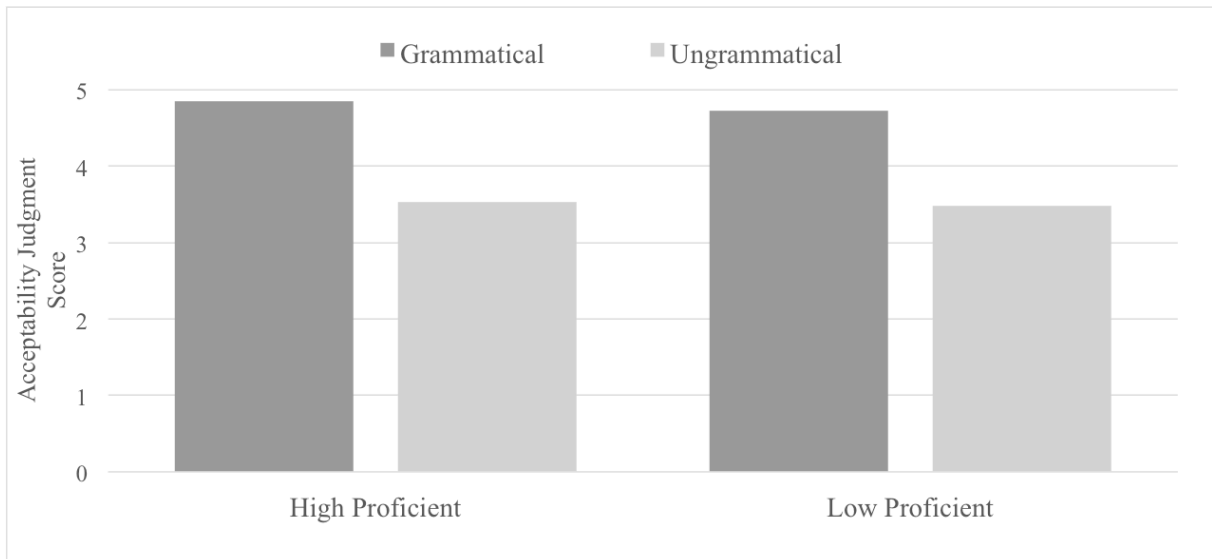
³⁸ Both grammatical and ungrammatical sentences have examples of the –ed and the –s morpheme.

means and medians of participants', high and low proficient, sentence score in grammatical and ungrammatical sentences can be observed below:

Table 5: Means and Medians of Participants' sentence score in the AJT task

Sentence type	Means	Medians
Grammatical high proficient	4.85	5
Ungrammatical high proficient	3.53	3.8
Grammatical low proficient	4,72	4.83
Ungrammatical low proficient	3.48	3.5

Figure 4: Grammatical and Ungrammatical Sentences Score Means of High and Low Proficient Participants in AJT



Subsequently, a normality test was conducted. The results displayed by the Shapiro-Wilk test indicate that the sentence score of the high proficient as well as the low proficient group, considering the comparison between grammatical and ungrammatical sentences, were not normally distributed:

Table 6: Shapiro-Wilk test for participants' sentence score distribution in AJT

Group	W	p value
High proficient group: grammatical and ungrammatical sentences	.667	2.40E-21
Low proficient group: grammatical and ungrammatical sentences	.717	8.54E-26

As a result, the Mann-Whitney was used to analyze this non-parametric data. The results suggest that the sentence score of the high proficient group – for the comparison between their grammatical and ungrammatical sentences – is statistically significant ($U= 2.5$, $Z=5.8766$, $p< .00001$). The same significance is found in the low proficiency group – for their grammatical and ungrammatical sentences comparison – ($U= 15$, $Z=4.02337$, $p< .00001$).

In what concerns the results of the sentence scores, it can be suggested that both high proficient and low proficient learners exhibit similar results to grammatical and ungrammatical sentences considering the morphemes under investigation. Therefore, it indicates that both the high proficient and the low proficient participants make use of their explicit and declarative knowledge, such as suggested by Souza & Oliveira (2017). As a consequence, it seems that it corroborates Jackendoff & Audring's (2016) proposal that defends that morphological schemas are part of the declarative knowledge of the language. Furthermore, Ullman (2001b) claims that there is productivity in the declarative memory of L2 speakers; thus, the rules may have been explicitly learned and accessed.

Unlike Silva (2016) who found significant difference between his high and low proficient groups, the results of this study suggest that both groups presented similar results in a speeded AJ task, which may be due to the fact that explicit knowledge was accessed. Besides, there could a different outcome if implicit or procedural knowledge was under investigation, such as the research conducted by Carneiro (2011). However, Carneiro (2011) found only a marginal difference between her high and low proficient participants' RT for the morpheme of past (-ed).

4.4 Experiment 2: Acceptability Judgment with Memory Load

4.4.1 Acceptability Judgment

Consistently with what was done in the first experiment, the results of both high and low proficient participants' RT are displayed, and after that, their sentence score in grammatical and ungrammatical sentences:

Table 7: Means and SD of Participants' RT in the AJWM task

Sentence type	Means	SD
Grammatical high proficient	6,370	2,710
Ungrammatical high proficient	6,780	2,910
Grammatical low proficient	8,270	3,250
Ungrammatical low proficient	8,180	3,650

As a consequence, the Shapiro-Wilk test was also administered to check if the participant's RT was normally distributed. Therefore, when the RT of both high and low proficient participants are contrasted in grammatical and ungrammatical sentences, the results suggest that they were not normally distributed:

Table 8: Shapiro-Wilk test for participants' RT distribution in AJWM

Group	W	p value
High proficient group: grammatical and ungrammatical sentences	.835	3.07E-15
Low proficient group: grammatical and ungrammatical sentences	.847	1.29E-11

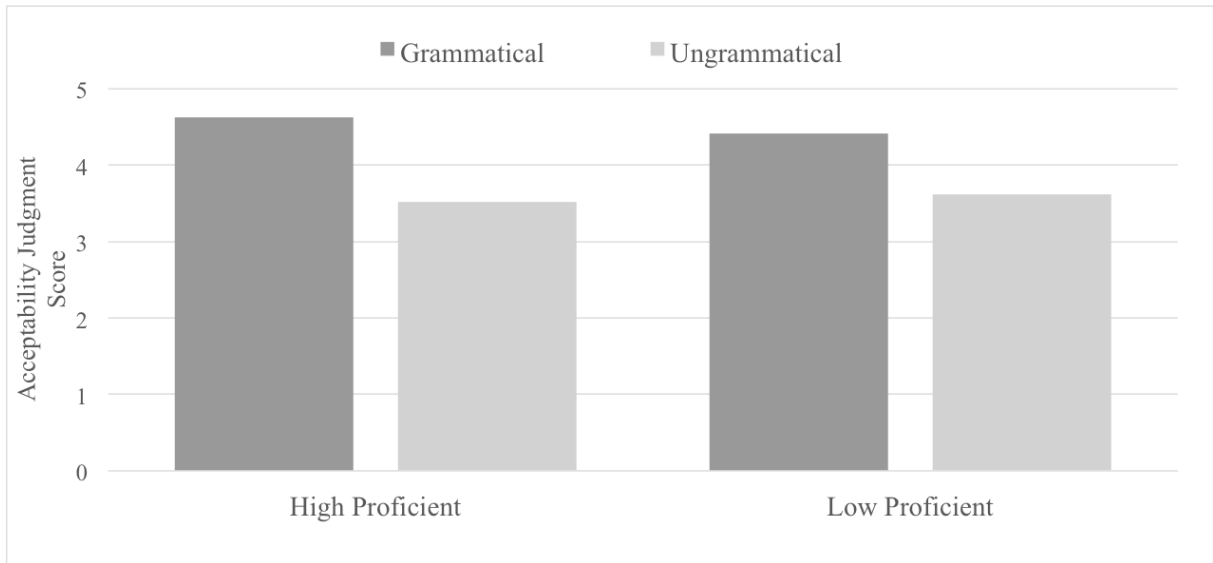
As a result, the Mann-Whitney test was conducted to study the data. Similarly to the results found in the first experiment, the RT of the high proficient group – contrasting their RT in grammatical and ungrammatical sentences – was not significant ($U= 257$, $Z= -0.6289$, $p> .5287$), along with the low proficient group – contrasting their RT in grammatical and ungrammatical sentences ($U= 107$, $Z= -0.20739$, $p> .83366$). The RT accounted here is for the whole experiment, namely, score and number memorization.

Accordingly, the same type of analysis carried out for the RT, was conducted for the sentence score. Means and medians of grammatical and ungrammatical sentences of both high proficient and low proficient participants are displayed as follows:

Table 9: Means and Medians of Participants' sentence score in the AJWM task

Sentence type	Means	Medians
Grammatical high proficient	4.65	4.64
Ungrammatical high proficient	3.52	3.66
Grammatical low proficient	4.41	4.57
Ungrammatical low proficient	3.62	3.66

Figure 5: Grammatical and Ungrammatical Sentences Score Means of High and Low Proficient Participants



Subsequently, the normality test Shapiro-Wilk was carried out in the sentences score of the participants. The results indicate that the sentence scores of the high proficient group – grammatical and ungrammatical sentences – and the low proficient group – grammatical and ungrammatical sentences – were not normally distributed:

Table 10: Shapiro-Wilk test for participants' sentences score distribution in AJWM

Group	W	p value
High proficient group: grammatical and ungrammatical sentences	.698	1.67E-20
Low proficient group: grammatical and ungrammatical sentences	.726	6.25E-16

Hence, the Mann-Whitney test was performed. When a comparison was drawn between the grammatical and ungrammatical sentence scores of the high proficient group, the results indicate that there is a statistical difference ($U= 51, Z=4.87655, p< .00001$), similar difference is found for the comparison of the grammatical and ungrammatical sentence scores of the low proficient group ($U= 57, Z=2.28129, p< .0226$).

Even though the sentence scores of the two groups did not have the same statistical reliability – such as in the first experiment – it is possible to observe that even with a memory load both groups presented results that show that proficiency did not affect memory capacity and processing. Such as discussed in the first experiment, the knowledge examined here seems to be the explicit and declarative knowledge, in accordance with Jackendoff & Audring (2016) and Souza & Oliveira (2017). Therefore, the participants were able to access their explicit knowledge since both high and low proficient participants displayed similar results in

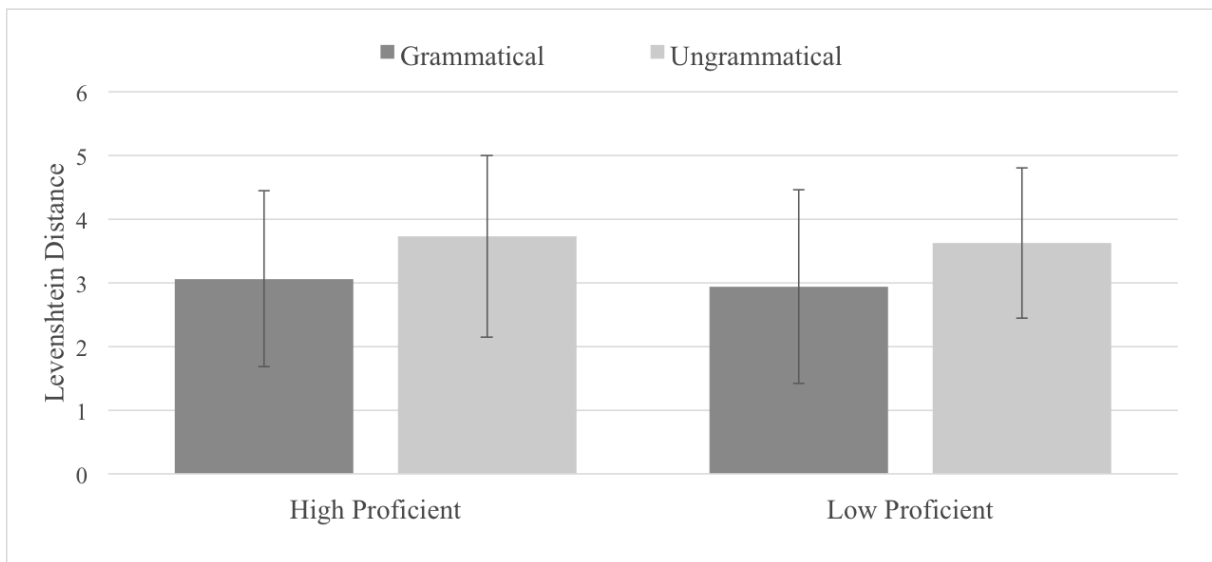
the sentence scores of grammatical and ungrammatical sentences – considering the morphemes that were investigated.

Despite the fact that Jesus (2012) found evidence proving that high proficient bilinguals presented more accuracy when declarative and procedural memory was being tested, the working memory capacity tested here seems to be a different construct. Therefore, the ungrammatical sentences of this study do not seem to influence high and low proficient participants' processing ability.

4.4.2 Memory Load

The Levenshtein Distance metric was applied in the memorized numbers of participants. The figure below exhibits how far the number that high and low proficient participants typed were from the correct ones. It is possible to observe that both high and low proficient participants had similar results for the numbers typed in grammatical and ungrammatical sentences:

Figure 6: High and Low Proficient Participants Means of Levenshtein Distance in AJWM



Similarly to the analysis carried out in the first part of this experiment, the same was done for the memory load that participants were supposed to memorize. As a consequence, after the Levenshtein metric was applied, the Shapiro-Wilk test was applied to check the normality of the data. The results indicate that the numbers typed in grammatical and ungrammatical sentences by the high and low proficient groups were normally distributed:

Table 11: Shapiro-Wilk test for participants' Levenshtein distance distribution in AJWM

Group	W	p value
High proficient group: grammatical and ungrammatical sentences	.968	.225
Low proficient group: grammatical and ungrammatical sentences	.982	.342

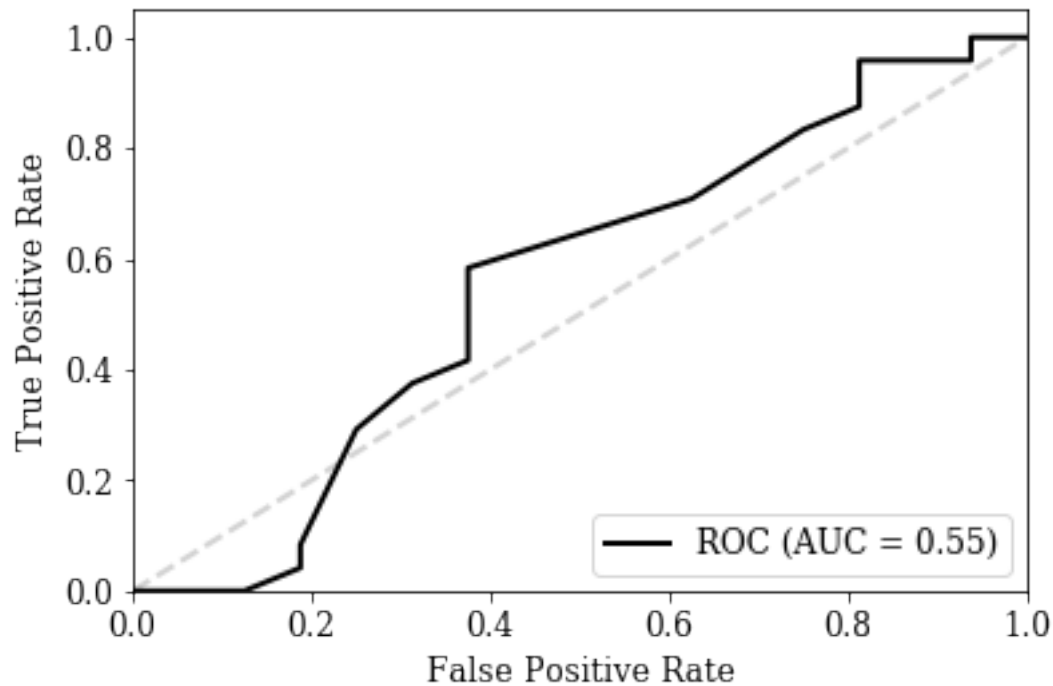
As the data was normally distributed, the unpaired *t* test was used to obtain information about the statistical relevance of the results. The results suggest that when the Levenshtein distance of the high proficient participants in grammatical ($M= 3.06$, $SD = 1.38$) and ungrammatical sentences ($M= 3.73$, $SD = 1.59$) was compared, there was no statistical significance ($p= .1267$, $t= 1.55$, $df= 46$, $SED = .431$). Similarly, low proficient learners presented no statistical significance in their results ($p= .1838$, $t= 1.36$, $df= 28$, $SED = .499$) when their performance in grammatical ($M= 2.94$, $SD = 1.52$) and ungrammatical ($M= 3.62$, $SD = 1.18$) sentences was contrasted.

As a result, the results indicate that the high proficient as well as low proficient participants did not have their WMC jeopardized by the grammaticality of the sentence. Thus, the ungrammatical sentences did not cause an extra processing load in comparison to the grammatical sentences. Furthermore, both of the groups exhibit similar results, as it is possible to see in Figure 6. This may be due to the fact that WMC cannot be regarded as part of the proficiency level since the two different groups' accomplishments are similar.

Although Finardi (2009) did not test the level of proficiency of her L2 participants, Azevedo (2012) conducted her research with low proficient learners, and both of them found corroborating evidence for L2 and WMC with their oral stimuli. As a consequence, WMC may not be affected by the level of proficiency, but by L2 per se. Furthermore, this study presented written stimuli for participants rather than oral stimuli, such as Finardi (2009) and Azevedo (2012) have done, which may have affected participants' performance.

In addition, the ROC curve analysis was applied to test the performance of high proficient – considering grammatical and ungrammatical sentences –, and low proficient participants – considering grammatical and ungrammatical sentences – with the result obtained in the Levenshtein Distance. Our goal was to check if high proficiency was connected to high WMC and if low proficiency was related to low WMC:

Figure 7: ROC curve of High and Low participants' performance in Memory Load in AJWM



The results from the ROC curve seem to indicate that the performance of high and low proficient participants overlap. As a result, it is not possible to predict that high proficient bilinguals will also have high WMC, and consequently, low proficiency does not entail low WMC.

4.5 Experiment 3: 2-back

The 2-back follows the same format as the other experiments. First, the analysis of the RT of participants is presented and after that the analysis of their performance. As this experiment concerned sequences of letters, there was no sentence performance to be contrasted; hence, the performance of high and low proficient participants is compared. Furthermore, this experiment takes into account the total amount of time spent to perform it, and not the amount of time spent in each item. As it is possible to observe in table 12, the high proficient participants needed more time to finish this experiment than the low proficient participants:

Table 12: Means and SD of participants' RT in the 2-back task

Group	Means	SD
High proficient	250.54*	66.45*
Low proficient	226.71*	51.53*

* Measure in seconds

The Shapiro-Wilk test was also conducted in this experiment since we aimed to see if the RT and performance were normally distributed among the high and low proficient participants. As it is possible to see from data below, the RT from both the high proficient group and the low proficient group was normally distributed. Furthermore, when the results from both the high and the low groups were compared, their distribution was normalized:

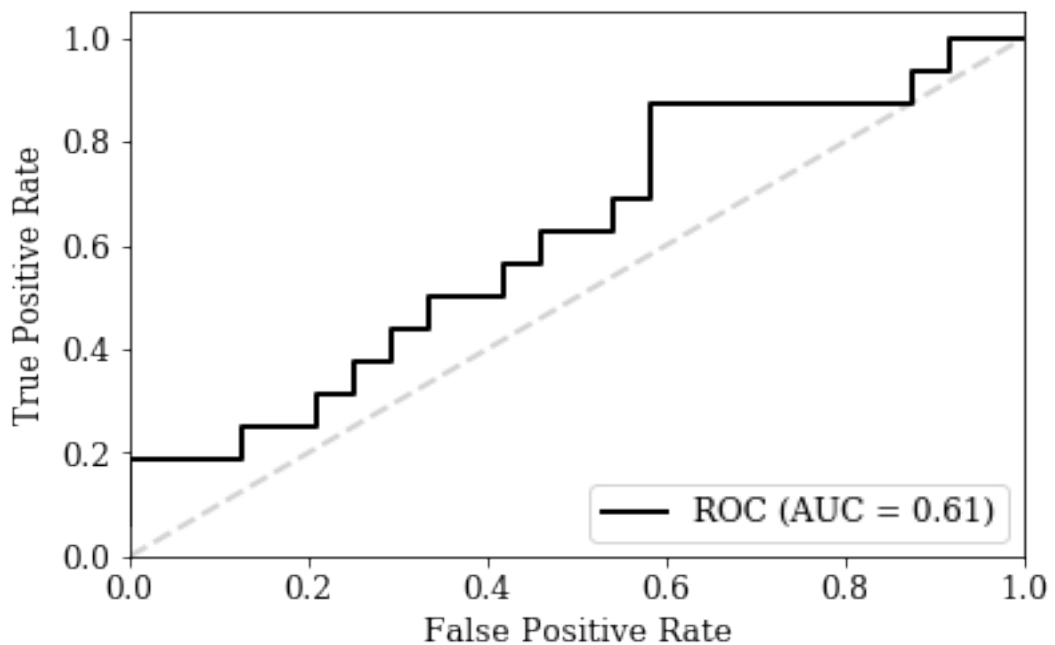
Table 13: Shapiro-Wilk test for participants' RT distribution in 2-back

Group	W	p value
High proficient group	.956	.367
Low proficient group	.955	.582
High vs. Low proficient groups	.962	.210

Since the RT of the two groups were normally distributed, we conducted an unpaired *t* test to analyze if there was any statistically significant difference when we compared the results of both high and low proficient learners. Despite the fact that the high proficient group was slower than the low proficient group, the results indicate that there was no statistically significant difference between their results ($p = .216$, $t = 1.25$, $df = 37$, $SED = 18.969$).

Subsequently, the ROC curve analysis was also conducted to compare the RT of the high proficient and low proficient participants. We aimed to investigate if WMC was correlated with proficiency:

Figure 8: ROC curve of High and Low participants' RT in 2-back



The results suggest that there is a poor chance of predicting if WMC and proficiency are related to the RT of the high and low proficiency participants investigated in this study, since the RT of the two groups almost overlap.

Accordingly, the performance – number of correct items, including targets and foils – of high and low proficient participants is displayed. Table 14 shows that the two groups' accuracy is very similar:

Table 14: Means and SD of Participants' performance in the 2-back task

Group	Means	SD
High proficient	179.58*	12.44*
Low proficient	178.47*	9.55*

*Number of correct items from 192

After that, the same analysis of normality, the Shapiro-Wilk test, was carried out to test if the performance of the groups was normally distributed:

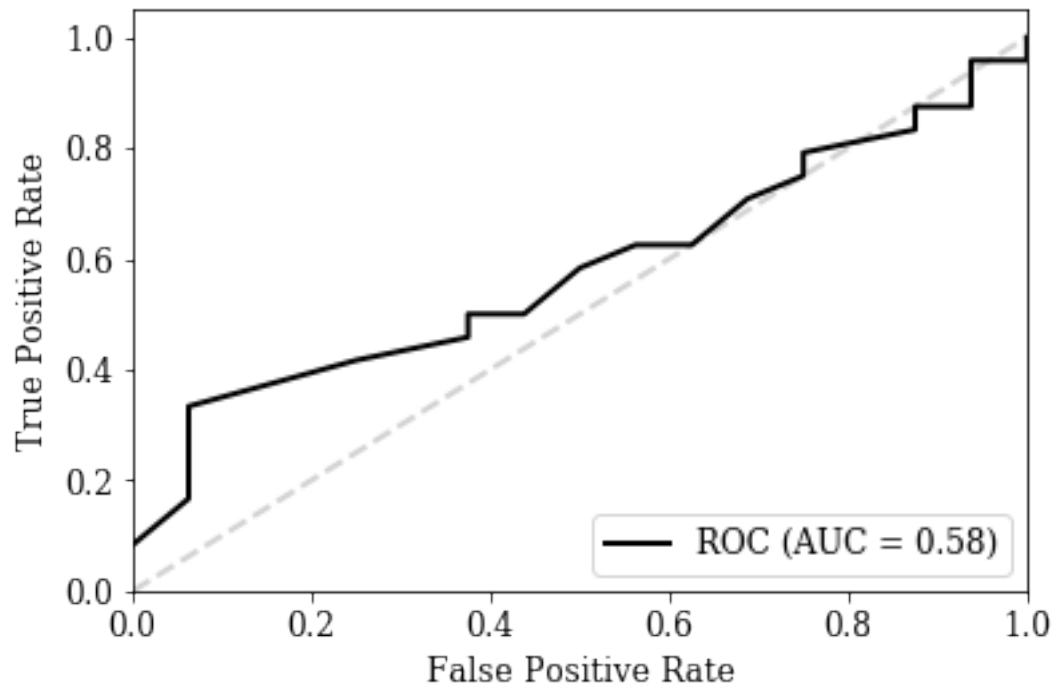
Table 15: Shapiro-Wilk test distribution for participants' performance in 2-back

Group	W	p value
High proficient group	.801	.0003
Low proficient group	.846	.012
High vs. Low proficient groups	.821	2.0E-05

As the data was not completely normally distributed for the participants' performance, the Mann-Whitney test was done. When we compared the performance of both high and low proficient participants, there was no significant difference in their results ($U= 147$, $Z=0.93819$, $p> .34722$).

Following the analysis, the ROC curve was also conducted for the performance of participants. Our goal was to investigate if proficiency and WMC were constructs that were related to the performance and level of proficiency of the participants:

Figure 9: ROC curve of High and Low participants' performance in 2-back



The results obtained indicate that the performance of both groups – high and low proficient – overlap, which means that their WMC cannot be predicted considering their level of proficiency.

As a consequence, both the RT and the performance of the high and low proficient groups in this experiment suggest that proficiency is not a predictor of working memory capacity since there was no significant difference between their results. Besides, as Finardi (2009) and Azevedo (2012) found results suggesting that L2 correlates with WMC, this may be due to the fact that bilingualism per se offers a cognitive advantage in comparison to monolinguals (BIALYSTOK *et al.*, 2009), which in turn could also be found in low proficient learners such as in Azevedo (2012).

CHAPTER 5

CONCLUSION

Concisely, this research aimed to investigate: (a) if WMC influences processing of the morphemes –ed and –s, and, (b) if the level of proficiency is correlated to the WMC.

In order to do so, data was gathered from Brazilian speakers of English as an L2, divided into high and low proficient with the Vocabulary Levels Test (VLT). The performance of the two groups was compared to check if they displayed statistically significant difference when grammatical and ungrammatical sentences were presented to them in two acceptability judgment tasks. One of the AJ tasks had a time restriction, in which participants had up to 6,000 milliseconds to rank each sentence, and the other one had an external memory load, as a consequence, participants were supposed to memorize a seven-digit number before ranking each sentence. The reaction time and score ranking – obtained through the Likert scale – of participants were compared. Therefore, the RT of the high proficient group was contrasted in grammatical and ungrammatical sentences, and the same was done for the low proficient group. The score ranking of grammatical and ungrammatical sentences was compared for the high proficient group, and for the low proficient group. The target sentences in the AJ tasks contained the inflectional morphemes third person singular (–s) and regular past verbs (–ed) in the grammatical sentences, and violations with absence of these morphemes in contexts where they are mandatory in the ungrammatical sentences.

Besides, the acceptability judgment task with external memory load and the 2-back task were administered so that we could have two types of tasks measuring WMC. As the 2-back task consisted in four lists containing sequences of letters – each list had 48 letters from which 40 were foils and 8 were targets –, the RT and performance of high and low proficient participants were compared. The performance of participants investigated in this task was the number of correct items that each participant obtained including targets and foils.

As a result, we did not find any effect that correlates WMC and processing of the morphemes –ed and –s as the participants' performance was not jeopardized by the grammaticality of sentence. Both high and low proficient participants could perceive violations of the morphemes under investigation. Moreover, the memory load of the AJ task did not affect the memory capacity and processing capacity of the two groups in grammatical and ungrammatical sentences, which may indicate that both groups accessed declarative and explicit knowledge (JACKENDOFF & AUDRING, 2016; SOUZA & OLIVEIRA, 2017). No effect was found proving the correlation between level of proficiency and WMC. The high

and low proficient groups exhibited similar results in both the AJWM and the 2-back tasks. This may be explained by the fact that bilingualism confers a cognitive advantage (BIALYSTOK *et al.*, 2009). This would explain how Azevedo (2012) found a correlation between WMC and bilingualism in low proficient participants.

5.1 Discussion

In the first experiment, we compared the performance of participants in a speeded acceptability judgment task. Participants were divided into high and low proficient, and grammatical and ungrammatical sentences regarding the selected inflectional morphemes – third person singular (–s) and regular past verbs (–ed) – were presented to them. We verified if there were statistically significant differences when we compared the mean RTs for judgment decisions to be made, and the mean judgment scores for both the grammatical and ungrammatical sentences presented for each group. After we conducted the statistical analysis, we discovered that there was no statistically significant difference between groups concerning their RTs, but the two proficiency groups exhibited statistically significant differences for their judgment score rankings when their performance in grammatical and ungrammatical sentences was compared.

The high proficient group presented statistically significant difference when their score ranking mean was compared in grammatical and ungrammatical sentences. Similarly, statistically significant difference was found for the comparison of the score ranking mean of grammatical and ungrammatical sentences of the low proficient group. This indicates that both groups of proficiency access their explicit and declarative knowledge despite the temporal restriction, as pointed out by Souza & Oliveira (2017). Hence, it seems to be in accordance with Jackendoff & Audring's (2016) proposal as they support the idea that morphological schemas exist as words, and therefore, are stored in declarative knowledge. Additionally, Ullman (2001b) argues that the declarative memory of L2 learners also displays productivity; notwithstanding, Ullman's (2001a, 2001b) proposal claims that mental lexicon and mental grammar are separate constructs.

In the second experiment, like in the first one, high and low proficient groups' RTs for judgment decision and judgment score ranking were compared regarding their performance in grammatical and ungrammatical sentences. There was no time restriction in this acceptability judgment task; instead, it had an external memory load. As a consequence, participants were supposed to memorize a random seven-digit-number between the sentence presentation and the judgment call. Even though this task had no time constraint, participants were instructed

to perform it as fast and accurate as possible. The sequence of numbers that participants typed was analyzed with the Levenshtein Distance metric, so that we would be able to check how different the sequences that the participants typed were from the correct ones. The same group of high proficient and low proficient participants performed this experimental task.

Such as in the speeded AJ task, no statistically significant difference was found for the RT of both high and low proficient learners when the RT mean difference of grammatical and ungrammatical sentences was compared within each group. However, we found statistical significance for the mean difference of score ranking of both high and low proficient groups in grammatical and ungrammatical sentences. As a result, the memory load did not impact the performance of high and low proficient participants in grammatical and ungrammatical sentences with the morphemes third person singular (-s) and regular past verbs (-ed). This may also be explained by the type of knowledge accessed by participants, which indicates that it is the explicit and declarative knowledge as suggested by Jackendoff & Audring (2016) and Souza & Oliveira (2017).

Notwithstanding, Jesus's (2012) results indicate that her high proficient group was more accurate when dealing with tasks that demanded that their declarative and procedural memory was controlled. However, the working memory capacity examined here seems to be a different process under investigation as Azevedo's (2012) low proficient participants exhibited correlation between L2 and WMC.

In what concerns the memory load of the second experiment, there was no statistically significant difference when the grammatical and ungrammatical sentences of both high and low proficient learners were compared. As a consequence, neither the high proficient nor the low proficient participants were affected by the grammaticality of the sentence; thus, the ungrammatical sentences were not more onerous to process than the grammatical ones. Accordingly, WMC was not influenced by proficiency level in this experiment as the high and low proficient group presented similar results. Azevedo's (2012) low proficient participants displayed significant results for the correlation of L2 and WMC. Besides, Finardi (2009) did not control the proficiency level of her participants, and her results corroborate the relation between L2 and WMC. Consequently, WMC may be related to L2 irrespectively of proficiency level, which may indicate that declarative and procedural memory may be influenced by the level of proficiency as suggested by Jesus (2012)

Furthermore, the ROC curve analysis contrasting the memory load results achieved by the high and low proficient group in grammatical and ungrammatical sentences revealed that the performance of the two groups overlap; therefore, it is not possible to draw a prediction

regarding proficiency level and working memory capacity.

In the third experiment, participants – the high as well as the low proficient group – were asked to carry out a 2-back task. The RT and performance of both groups were compared. In this experiment, the RT and performance – number of correct items – of high and low proficient participants are contrasted, as there were no experimental sentences. The results reveal that there was no statistical significant difference when both the RT and the performance of high and low proficient learners were compared.

The ROC curve analysis of high and low proficient participants' RT in this experiment shows that there is a poor chance of predicting if proficiency and WMC are linked. Besides, the ROC curve analysis, conducted for the performance of the two groups of participants, displays that their performance almost overlaps. As a result, it also suggests that the WMC of participants is not related to their proficiency.

In addition, the results found by Finardi (2009) and Azevedo (2012) indicate that L2 is associated with WMC; therefore, it can mean that bilingualism confers a cognitive advantage in relation to monolinguals, such as proposed by Bialystok *et al.* (2009). This can explain why Azevedo (2012) found the correlation between WMC and her participants, who were a group of low proficient learners of English.

5.2 Limitations of the Study

The acceptability judgment tasks' designs had fewer ungrammatical sentences than grammatical sentences. Hence, future studies investigating the same topic could have more ungrammatical sentences in order to check if the results discussed here were due to the experimental design, or if it is also possible to corroborate these results when there are more target ungrammatical sentences.

Furthermore, all participants performed the three tasks at once, and later on the VLT. It is important to take into consideration that the three experiments combined lasted from 35 minutes to an hour, depending on the participant. The two experiments concerning WMC were the last ones that the participants carried out. The AJWM was the one that took more time to be concluded. As a result, when participants performed the 2-back task, participants could have been weary. Future studies may test the effect of performing fewer tasks so that the performance of participants may not be affected.

Additionally, as the working memory capacity tasks were the last ones to be conducted in this study, future studies could administer them previously along with the proficiency level test, so that participants could be selected based not only on their

proficiency level, but also on their WMC.

This study compared participants' score means and RT means in AJT and AJWM tasks considering the target grammatical and ungrammatical sentences. Future studies could compare the sentences score means and RT means in similar tasks with the similar amount of characters, but also with the same level of complexity as the sentences varied in complexity.

All the stimuli presented in this research to participants – including sentences, numbers and letters – were visually presented while Finardi (2009) and Azevedo's (2012) experiments were orally presented to participants. Besides, the grammatically judgment tasks investigated by McDonald (2006, 2008b) were orally presented to participants. Therefore, future studies trying to correlate WMC and proficiency could compare the performance of participants in experiments visually and orally presented, so that the different experimental paradigm results could we studied.

Moreover, an experiment measuring implicit knowledge, such as done by Carneiro (2011), could be carried out to check if the performance of high and low proficient participants presents a different outcome when contrasted to WMC.

REFERENCES

- ALBERT, M. S. & HEATON, R. K. Intelligence testing. In: ALBERT, M. S. & MOSS, M. B. (Eds). *Geriatric Neuropsychology*. New York: Guilford Press, p. 13-32, 1988.
- ALDERNSON, J. C. *Diagnosing Foreign Language Proficiency: The Interface between Learning and Assessment*. London/New York: Continuum, 2005.
- AZEVEDO, A. F. Listening comprehension and working memory capacity in beginning L2 learners: an exploratory study. 2012. 186 f. Mestrado em Linguística e Letras. Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2012. .
- BAAAYEN, R. H., DIJKSTRA, T., & SCHREUDER, R. Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, v. 37, p. 94–117, 1997.
- BADDELEY, A. D. Working memory. *Science*, v. 255, p. 556-559, 1992b.
- _____. Working Memory in Second Language. In: WEN, Z. E.; MOTA, M. B. & MCNEIL, A. (Eds) *Working Memory in Second Language Acquisition and Processing*, p. 17-28, 2015.
- BARD, E. G.; ROBERTSON, D. & SORACE, A. Magnitude Estimation of Linguistic Acceptability. *Language*, v. 72, n. 1, p. 32-68, 1996.
- BIALYSTOK, E.; CRAIK, F.I.M.; GREEN, D. W. & GOLLAN, T.H. Bilingual minds. *Psychological Science*, v. 10, n.3, p. 89-129, 2009.
- BIRDSONG, D. *Metalinguistic performance and interlinguistic competence*. New York: Springer-Verlag, 1989.
- BOOJI, G. *Construction morphology*. Oxford: Oxford University Press, 2010.
- _____. Inheritance and motivation in Construction Morphology. In: GISBORNE & HIPPISEY (eds.), *Defaults in morphological theory*. Oxford: Oxford University Press, p. 18-39, 2017.
- CARNEIRO, M. M. *Morfologia de flexão verbal no inglês como L2: uma abordagem a partir da Morfologia Distribuída*. 2008. 149f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.
- _____. *Processamento Linguístico de Marcas de Morfologia de Flexão em Contexto de Inglês como Segunda Língua*. 2011. 148f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.
- CHOMSKY, N. *Syntactic Structures*. The Hague: Mouton. 1957.
- _____. A minimalist program for linguistic theory. In: K. Hale and S. J. Keyser (eds.). *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, Cambridge, Mass.: MIT Press, p. 1-52, 1993
- _____. *The minimalist program*. Cambridge, London: MIT, 420f, 1995.

CLAHSEN, H. & FELSER, C. Grammatical processing in language learners. *Applied Psycholinguistics*, v. 27, p. 3–42, 2006a.

CLAHSEN, H. & FELSER, C. How native-like is non-native language processing? *Trends in Cognitive Sciences*, v. 10, p. 564–570, 2006b.

COWAN, N. *Working memory capacity*. Hove, East Sussex, England: Psychology Press, 2005.

_____. The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, v. 19, p. 51–57, 2010.

CROMHEEKE, S., & MUELLER, S. C. The power of a smile: stronger working memory effects for happy faces in adolescents compared to adults. *Cogn. Emot.* v. 30, p. 288–301, 2016. doi: 10.1080/02699931.2014.997196

ELLIS, R. *et al. Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol, UK: Multilingual Matters, 391f, 2009.

ELLIS, N. C. At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, v. 27, p. 305–352, 2005.

ENGLE, R. W., & KANE, M. J. Executive attention, working memory capacity, and a two-factor theory of cognitive control. In: B. ROSS (Ed.), *The psychology of learning and motivation*. New York: Academic Press, p. 145–199, 2004.

FERNÁNDEZ, E. M.; SOUZA, R. A.; CARANDO, A. Bilingual innovations: Experimental evidence offers clues regarding the psycholinguistics of language change. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 20, n. 2, p. 251–268, 2017.

FINARDI, K. R. Working Memory Capacity and the Acquisition of a Syntactic /structure in L2 Speech. 2009. 231 f. Doutorado em Letras (Inglês e Literatura Correspondente). UFSC, Florianópolis. 2009.

GATHERCOLE, S.E. Working memory and language. In: GASKELL, M.G. (Ed.), *The Oxford handbook of psycholinguistics*. Oxford: Oxford University Press, p. 757–769, 2007.

GRAY, J. R.; CHABRIS, C. F. & BRAVER, T. Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, v. 6, p. 316–322, 2003.

HAGOORT, P. On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, v. 9, p. 416–423, 2005. doi: 10.1016/j.tics.2005.07.004

HALLE, M. & MARANTZ, A. Distributed Morphology and the pieces of inflection. In: HALE, K.; KEYSER, S.J. (Ed.). *The View from Building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, p. 111–176, 1993.

HAN, Z-H. Grammatical morpheme inadequacy as a function of linguistic relativity: A longitudinal study. In: HAN, Z.-H. & CADIerno, T. (Eds.), p. 154–182, 2010.

_____. Forty years later: Updating the Fossilization Hypothesis. *Language Teaching*,

v. 46, n. 2, p. 133-171, 2013.

HEBB, D.O. *The Organization of Behavior*. New York: Wiley & Sons, 1949.

JACKENDOFF, R. Semantic and morphological regularities in the lexicon. *Language*, v. 51, p. 639-671, 1975.

_____. *The architecture of the language faculty*. Cambridge Mass.: MIT Press. 1997.

JACKENDOFF, R. & AUDRING, J. Morphological schemas: theoretical and psycholinguistic issues. *The Mental Lexicon*. v. 11, n. 4, p. 467–493, 2016.

JESUS, D. B. The effect of L2 proficiency on the declarative and procedural memory systems of bilinguals: a psycholinguistic study. 2012. 108 f. Mestrado em Letras (Inglês e Literatura Correspondente). UFSC, Florianópolis, 2012.

JIANG, N. Morphological insensitivity in second language processing. *Applied Psycholinguistics*, v. 25, p. 603-634, 2004b.

_____. Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, v. 57, n. 1, p. 1-33, 2007.

JUST, M. A.; CARPENTER, P. A. & WOOLLEY, J. D. Paradigms and processes and in reading comprehension. *Journal of Experimental Psychology: General*, v. 3, n. 2, p. 228- 238, 1982.

KANE, M. J.; CONWAY, A. R. A.; MIURA, T. K. & COLFLECH, G. J. H. Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, v. 33, p. 615-622, 2007.

KOPF, J.; DRESLER, T., REICHERTS, P.; HERMANN, M. J. & REIF, A. The effect of emotional content on brain activation and the late positive potential in a word n-back task. *PLoS ONE*. 2013. 8:e75598. doi: 10.1371/journal.pone.0075598

LANGACKER, R. *Foundations of Cognitive Grammar*. Stanford, CA: Stanford University Press. v. 1, 1987.

LENNEBERG, E. *Biological foundations of language*. New York: Wiley, 1967.

LEVELT, W.; ROELOFS, A. & MEYER, A. A theory of lexical access in speech production. *Behavioral and Brain Sciences*, v. 22, p. 1–75, 1999.

LEVENSHTEIN, V. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physice-Doklady*, V. 10, p.707-710, 1966.

LIBBEN, G. Why study compounds? An overview of the issues. In G. Libben & G. Jarema (Eds.), *The representation and processing of compound words*. Oxford: Oxford University Press, p. 1–21, 2006.

MARINIS, T. Using on-line processing methods in language acquisition research. In: UNSWORTH, S. & BLOM, E. (Eds.), *Experimental Methods in Language Acquisition*

Research. John Benjamins, Amsterdam, p. 139-162, 2010.

MARX, I.; DOMES, G.; HAVENSTEIN, C.; BERGER, C.; SCHULZE, L. & HERPERTZ, S. C. Enhanced emotional interference on working memory performance in adults with ADHD. *World J. Biol. Psychiatry*. v. 12, p. 70–75, 2011. doi: 10.3109/15622975.2011.599213

MAY, C. P.; HASHER, L., & KANE, M. J. The role of interference in memory span. *Memory & Cognition*, v. 27, p. 759–767, 1999.

MCDONALD, J. L. Beyond the critical period: Processing- based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory & Language*, v. 55, p. 381-401, 2006.

_____. Differences in the cognitive demands of word order, plurals, and subject-verb agreement constructions. *Psychonomic Bulletin & Review*. v. 15, p. 980-984, 2008b.

MITCHELL, D. On-line methods in language processing: introduction and historical review. In: CARREIRAS, M; CLIFTON JR, C. (Ed.). *The on-line study of sentence comprehension: eye-tracking, ERPs and beyond*. New York: Psychology Press. p. 16-32, 2004.

_____. *Teaching and Learning Vocabulary*. Boston, MA: Heinle & Heinle, 1990.

OBERAUER, K. Binding and inhibition in working memory— Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, v. 134, p. 368–387, 2005.

OLIVEIRA, C. S. Processing, Representation and Learnability of the Resultative Construction by Brazilian Portuguese-English Bilinguals. 2016, 202f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2017.

PEIRCE, J. W. Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatic*, v. 2, n. 10, p. 1-8, 2009.

ROBERTS, R. & GIBSON, E. Individual differences in sentence memory. *Journal of Psycholinguistic Research*, v. 31, p. 573–598, 2002.

RONNLUND M.; NYBERG L.; BACKMAN L. & NILSSON L. G. Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population- based study. *Psychology and Aging*. v. 20, p. 3–18, 2005. [PubMed: 15769210]

SALTHOUSE, T. A. When does age-related cognitive decline begin? *Neurobiol Aging*. v. 30, n. 4, p. 507–514, 2009. doi:10.1016/j.neurobiolaging.2008.09.023.

SALTHOUSE, T. A.; BABCOCK, R. L. Decomposing adult age differences in working memory. *Developmental Psychology*, v. 27, p. 763-776, 1991.

SCHIMITT, N. Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, v. 64, n. 4, p. 913–951, 2014.

SCHUTZE, C. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press, 1996.

SILVA, J. S. Exploring a vocabulary test and a judgment task as diagnoses of early and late bilinguals' L2 proficiency. 2016. 150f. Tese (Doutorado) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte. 2016.

SLABAKOVA, R. The bottleneck of second language acquisition. *Foreign Language Teaching and Research*, v. 46, n. 4, p. 543-559, 2014.

SORACE, A. Possible manifestations of shallow processing in advanced second language speakers. *Applied Psycholinguistics*, v. 27, n. 1, p. 88–91, 2006b.

_____. Pinning down the concept of 'interface' in bilingualism. *Linguistic Approaches to Bilingualism*, v. 1, n.1, p. 1–33, 2011.

SOUZA, R. A. Déficit representacional entre falantes de L2? Uma reflexão em torno de dois modelos da arquitetura global da gramática. *Signótica*, v.22, n.2, 2010.

_____. Memorização de sentenças com construção específica de uma das línguas de bilíngues. *Revista Letrônica*, Porto Alegre, Pontifícia Universidade Católica-RS, v. 7, n. 1, p. 86-112, 2014.

SOUZA, R. A.; DUARTE, D. & BERG, I. B. A avaliação de habilidades linguísticas em L2: Uma questão metodológica em estudos de influências translinguísticas. IN: BUCHWEITZ, A.; MOTA, M. B. (orgs.). *Linguagem e Cognição: Processamento, Aquisição e Cérebro*. Porto Alegre: EDIPUCRS, p. 267-292, 2015.

SOUZA, R. A. & OLIVEIRA, C. A. The learnability of the resultative construction in English L2: a comparative study of two forms of the AJT. *Revista da Abralín*, v. 8, n. 2, p. 375-410, 2014.

_____. Are bilingualism effects on the L1 byproducts of implicit knowledge? Evidence from two experimental tasks. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 25, n.3, p.1685-1716, 2017

SOUZA, R. A.; OLIVEIRA, C. S.; SOARES-SILVA, J.; PENZIN, A. & Santos, A. A. Estudo sobre um parâmetro de tarefa e um parâmetro amostral para experimentos com julgamentos de aceitabilidade temporalizados. *Revista de Estudos da Linguagem*, v. 23, n. 1, 2015.

SOUZA, R. A; SILVA, J. S. Exploring the measurement of vocabulary size to differentiate Brazilian Portuguese-English bilinguals' access to grammatical knowledge in the L2. *Revista Linguística*, v. 11, n. 1, p. 187-204, 2015.

SOUZA, R.; SOARES-SILVA, J. & SILVA, G. Brazilian Portuguese-English bilingualism does not affect metalinguistic awareness of L1 constraints in two argument structure constructions. *Ilha do Desterro*, Florianópolis, UFSC, v. 69, n. 1, p. 17-32, 2016.

ULLMAN, M. T. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, v. 30, n. 1, p. 37-69, 2001a.

_____. The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition*, v. 4, p. 105-122, 2001b.

WHITE, L. & GENESEE, F. How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, v. 12, p. 233–265, 1996.

APPENDIX**APPENDIX 1- Vocabulary Levels Test (VLT)****Instructions**

This is a vocabulary test. You must choose the right word to go with each meaning. Choose the words on the right to the meanings on the left. Here is an example:

- | | |
|-------------------------------|--|
| 1. part of a house | a) business
b) clock
c) horse
d) pencil
e) shoe
f) wall |
| 2. animal with four legs | a) business
b) clock
c) horse
d) pencil
e) shoe
f) wall |
| 3. something used for writing | a) business
b) clock
c) horse
d) pencil
e) shoe
f) wall |

The correct answers are: 1. (part of a house) and **f** (wall); 2. (animal with four legs) and **c** (horse); 3. (something used for writing) and **d** (pencil).

APPENDIX 2 – Acceptability Judgment with Time Restriction

Training sentences

My mother is taller than my father.
 I will looked the television last week.
 Brian come home at 7.
 This exam is very easy.
 I love playing tennis.
 I would marry you.

Target sentences

Man land on the moon in 1969.
 Robert phoned the police yesterday.
 Mary wash her clothes yesterday.
 The maid slowly cleaned the floor.
 The gardener planted some rose trees.
 He love getting up at 10 o'clock.
 My uncle eats breakfast every morning.
 John often kisses his wife.
 Mary usually play with her dog.
 He usually buys food from the market.

Distractor sentences

Dogs are more intelligent than cats.
 My teacher is usually late for class.
 Sarah has a headache yesterday morning.
 The two dogs was brown.
 David is in his thirties.
 I would rather just water than coke.
 How come did they break up?
 You have better leave the room now.
 The kids will play, no matter the rain.
 Despite the food, you have to eat.
 It doesn't matter what, you have to eat.
 I must eat fruits, no matter the taste.
 It's not important who did she talk to.
 Whatever the secret, I won't tell.
 She better had listen to her mother.
 I gave the library those books.
 We were washing the car for ours father.
 Mary can't play because she is studying.
 There was a picture on the wall.
 The singer went to the concert.
 I bought two new chairs for my room.
 Brian has not visited the museum.

Ten years ago Jackie bought a house.

Jeff is an excellent student.

I went on a fishing trip last summer.

Tom is watching a good film on TV.

The baby was crying since he was hungry.

My husband buys dinner for us.

I was in London last weekend.

Martha has two sisters and six cousins.

APPENDIX 3 – Acceptability Judgment with Memory Load

Training sentences

I didn't waited for him.	9780920
Moses imagined to whom what he said.	3485925
We usually buy our food from the local market.	2527606
Maria began to play tennis in 2015.	3089990
I never buy live albums.	2589948
The farmer walked his horse to the river.	6596186
They often are too tired.	3935482
You should to wear formal clothes.	6853635
It was a extremely frightened scene.	0759532
That could be a little dangerous.	5681423

Target sentences

The daughter stuff the turkey in an hour.	0698791
The army battalion rushed to the battle.	2622887
The family prospered in their new house.	1310621
The balloon pop with a loud bang.	4628237
The lion attacked the sickly deer.	9316005
A hairdryer blows out hot air.	8931477
The waitress tastes the dessert behind the door.	2271824
The politician often thank the voters.	7835837
The deer often strips the forest of its leaves.	9892963
The hunter often tracks the tiger across the valley.	1987162

Distractor sentences

I would rather a documentary than fiction.	4422104
Brazilian beaches are famous around the world.	1716396
The game will happen, no matter the weather.	8750452
A handler is a person who trains dogs.	8009236
Fraternity parties are often very noisy.	6947094
Whatever the salary, I won't take the job.	4039972
I was so tired after that walk.	9043767
Dimes and quarters are needed for parking meters.	7352108
How come did you go to that party?	1792481
The universities want to recruit more students.	9208948
Despite what does she say, we will talk to John.	3971812
I have had enough of the neighbor's noise.	0108729
We sent the box the envelopes.	0493085
She is proud of her achievements.	0188435
It doesn't matter the subject, you have study.	5789759
You were supposed to being here at time.	6196553
You have better study hard for this test.	0878776

Jumbo jets are more comfortable than buses.	2385170
Shopping malls have become increasingly popular.	3181851
You should see the doctor, no matter what the cost.	5081939
A clairvoyant is a person who sees the future.	6827209
A gnu is an animal which lives in the Savannahs.	7876776
It's not important what does Bob work with.	2675365
I can't wait to see you again.	2644094
The airplane has landed at the airport.	4953004
My teacher aren't very interested in her work.	0321926
You better had leave your wallet in the apartment.	6767317
We is getting a friend to repair the roof.	7049461
Two men is arrested yesterday at the mall.	8413597
I were born in Mexico in 1990.	0170818

APPENDIX 4 – 2-back

Training

List A

q-F-B-r-X-X-x-M-m-K-b-B-m-Q-M-X-h-B-H-x-K-q-F-f-F-k-K-m-R-h-H-m-B-r-B-F-q-H-Q-r-F-R-h-K-x-k-R-q

List B

R-q-H-k-F-f-R-b-B-b-F-m-K-h-X-b-X-h-Q-h-F-k-Q-q-Q-k-M-k-R-x-R-b-M-h-M-r-R-f-X-f-B-h-K-m-M-q-X-x

List C

F-x-H-m-F-x-X-m-H-f-Q-r-Q-b-B-m-X-m-F-h-F-k-M-h-H-h-B-q-B-k-K-k-R-b-R-x-Q-x-M-k-R-r-F-q-Q-k-R-b

Target

List D

k-F-x-B-r-H-q-Q-k-F-k-M-r-R-r-X-b-X-q-R-Q-K-k-X-r-B-h-B-F-f-H-b-H-m-M-m-Q-f-X-F-b-H-h-M-k-Q-x-M

List E

F-m-Q-h-B-r-b-F-m-F-x-R-r-f-X-b-x-Q-K-k-H-q-B-q-K-x-K-q-Q-r-M-r-H-h-H-b-B-k-F-x-M-m-M-r-X-f-H-k

List F

Q-h-K-m-Q-q-F-k-F-x-X-m-R-M-h-B-h-M-k-K-k-b-R-b-M-x-Q-x-R-b-H-h-X-b-F-q-H-q-F-f-B-x-K-r-R-r-M-f

List G

R-b-Q-f-X-x-X-k-B-k-X-h-R-h-F-m-F-h-B-b-M-h-m-Q-f-F-k-M-q-Q-q-R-x-K-k-R-h-r-M-m-X-b-Q-b-H-k-F-r