UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Letras
Programa de Pós-Graduação em Estudos Linguísticos

Bruna Rodrigues Fontoura

Functional Morphology and Working Memory: a sentence recall investigation with Brazilian Portuguese-English Bilinguals

Belo Horizonte

2022

Bruna Rodrigues Fontoura

Functional Morphology and Working Memory: a sentence recall investigation with Brazilian Portuguese-English Bilinguals

Tese 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 Doutor 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

Co-orientador: Prof. Dr. Cândido Samuel Fonseca de Oliveira

Belo Horizonte

2022

F684f

Fontoura, Bruna Rodrigues.
Functional Morphology and Working Memory [manuscrito] : a sentence recall investigation with Brazilian Portuguese-English Bilinguals / Bruna Rodrigues Fontoura. – 2022.
1 recurso online (177 f. : il., tabs., grafs., p&b., color.) : pdf.

Orientador: Ricardo Augusto de Souza.

Coorientador: Cândido Samuel Fonseca de Oliveira.

Área de concentração: Linguística Teórica e Descritiva.

Linha de Pesquisa: Processamento da Linguagem.

Tese (doutorado) – Universidade Federal de Minas Gerais,

Faculdade de Letras.

Bibliografia: f. 126-131.

Apêndices: f. 132-177.

Exigências do sistema: Adobe Acrobat Reader.

1. Língua inglesa – Morfologia – Teses. 2. Língua inglesa – Flexão – Teses. 3. Língua inglesa – Estudo e ensino – Falantes estrangeiros – Teses. 4. Aquisição da segunda linguagem – Teses. I. Souza, Ricardo Augusto de. II. Oliveira, Cândido Samuel Fonseca de. III. Universidade Federal de Minas Gerais. Faculdade de Letras. IV. Título.

CDD: 420.7

Ficha catalográfica elaborada pelo Bibliotecário Israel José da Silva – CRB/6-2128

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

FOLHA DE APROVAÇÃO

Functional Morphology and Working Memory: a sentence recall investigation with Brazilian Portuguese-English Bilinguals

BRUNA RODRIGUES FONTOURA

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

Aprovada em 04 de abril de 2022, pela banca constituída pelos membros:

Prof(a). Ricardo Augusto de Souza - Orientador

UFMG

Prof(a). Cândido Samuel Fonseca de Oliveira - Coorientador

CEFET-MG

Prof(a). Mailce Borges Mota

UFSC

Prof(a). Amanda Post da Silveira

UFJ

Prof(a). Deise Prina Dutra

UFMG

Prof(a). Marisa Mendonça Carneiro

UFMG

Belo Horizonte, 04 de abril de 2022.

Documento assinado eletronicamente por **Deise Prina Dutra**, **Professora do Magistério Superior**, em 05/04/2022, às 14:14, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Marisa Mendonca Carneiro**, **Professora do Magistério Superior**, em 05/04/2022, às 15:11, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Amanda Post da Silveira**, **Usuário Externo**, em 06/04/2022, às 16:52, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Cândido Samuel Fonseca de Oliveira**, **Usuário Externo**, em 06/04/2022, às 18:37, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Mailce Borges Mota**, **Usuário Externo**, em 08/04/2022, às 16:03, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Ricardo Augusto de Souza**, **Servidor(a)**, em 13/04/2022, às 15:13, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1324603** e o código CRC **A3AF52E6**.

To everyone that made it possible

ACKNOWLEDGEMENTS

This Ph.D. wouldn't have been possible without the guidance of my advisor Dr. Ricardo de Souza. I have learned from you since the undergraduate course and your teachings will never cease to inspire me.

I would like to express my gratitude to my co-advisor Dr. Cândido de Oliveira for accepting this challenge. Thank you for your kindness, patience, and great contributions to this study.

Thanks to Dr. Mailce Mota, Dr. Marisa Carneiro, Dr. Deise Dutra, and Dr. Amanda Post for the wonderful suggestions and comments on the study.

I am grateful to Dr. Mara Guimarães and her students Gabriela and Maria Clara for helping with the sentence ratings. Thanks for the hard work.

There would be no task if Weslley hadn't helped with the script. Thanks a lot!

I found a lot of support over the years along with the girls of the Psycholinguistic Lab. Special thanks to Clarice, Flavia, Isabelle, Jamila, Kelly, and Thaís (who deserves a lot of recognition).

Thanks to all the participants that kindly took part in this study.

The professors I had at UFMG helped to shape me and I see myself doing many things that you have taught me. I will keep you forever in my thoughts.

Thanks to UFMG for this amazing journey. All the opportunities made me a better person and professional.

I have had the privilege to keep learning so much from my students and my colleagues at UEMG-Passos. I am very honored for this opportunity.

I could not forget to thank my parents, Flávio and Márcia, who have believed me since the beginning of this journey, from my undergraduate course to the Ph.D.

I'm very thankful to my grandma Odracy and my aunt Flávia for dreaming with me during all these years. I am sure grandpa Antônio Carlos is still dreaming with us.

Thanks to my brother Flávio.

Thanks to my beloved cats Elis, Nina, and Branca. Your love has brightened up my days.

Thanks to all scientists worldwide who haven't given up despite all the adversities in these last years. Thanks for the privilege of being alive and vaccinated.

When I find myself in times of trouble [...]

There will be an answer, let it be

(The Beatles)

ABSTRACT

This study concerns functional morphology violation detection and correction by English second language (L2) learners – higher proficiency and lower proficiency – and English natives. We investigated whether proficiency influenced the three groups' mastery to produce the 3rd person singular morpheme (-s) in sentences with long-distance agreement (e.g. The engineer with the tools always builds big houses) and local-distance agreement (e.g. He always drinks orange juice) where the morpheme was missing. Two of our target conditions have long-distance ungrammatical sentences and two have local-distance ungrammatical sentences that participants were expected to correct, and our control condition has grammatical sentences that had to be reproduced. In addition, we analyzed whether these groups' working memory capacity had a role when they reproduced grammatical sentences and judged sequences of images and letters. In order to accomplish these objectives, we carried out a Sentence Recall Task (SRT) with an intermediary memory task in the format of letters and images that participants had to judge, and their proficiency was measured through the Vocabulary Levels Test (VLT). We gathered data from twenty higher proficiency participants, twenty-one lower proficiency participants, and twelve native participants. The data from our task suggest that proficiency was a determinant factor in predicting the L2 learners' ability to correct long-distance and local-distance agreement ungrammatical sentences because the higher proficiency group outperformed the lower proficiency group in all target conditions. However, surprisingly, the native group showed no significantly different performance from the lower proficiency group in one of our long-distance conditions and one of our local-distance conditions. There was only a marginal difference between the performance of higher proficiency and lower proficiency groups in reproducing grammatical sentences, but we detected a significant difference between the performance of the native group and the bilingual groups. We also found meaningful differences between the groups in images and letters judgment, as natives were statistically better than the bilinguals groups, and the higher proficiency group was significantly different from the lower proficiency group. The results suggest that proficiency influenced the L2 learners' ability to detect and correct ungrammatical sentences, and working memory capacity had an effect when participants reproduced grammatical sentences and judged the letters and images. We discuss our findings with the Bottleneck Hypothesis (SLABAKOVA, 2013, 2014), the Lexical Development in the L2 (JIANG, 2000), the Relational Morphology (RM) (JACKENDOFF & AUDRING, 2016), and the Phonological/Executive Model (P/E Model) (WEN; MOTA; MCNEILL, 2013).

KEYWORDS: L2; Functional Morphology; Working Memory; Sentence Recall Task.

RESUMO

Este estudo trata da detecção e da correção de violação de morfologia funcional por aprendizes de segunda língua – maior proficiência e menor proficiência – e nativos do inglês. Investigamos se a proficiência influenciou a habilidade dos três grupos de produzir o morfema de 3ª pessoa do singular do inglês (-s) em sentenças com concordância de longa distância (ex. *The engineer with the tools always builds big houses)* e de curta distância (ex. *He always drinks orange juice*) onde faltava o morfema. Duas de nossas condições alvo têm sentenças agramaticais de longa distância e duas têm sentenças agramaticais de curta distância que os participantes deveriam corrigir, e nossa condição controle tem sentenças gramaticais que precisavam ser reproduzidas. Além disso, analisamos se a capacidade de memória de trabalho desses grupos teve papel quando eles reproduziram frases gramaticais e julgaram sequências de imagens e letras. Para cumprir esses objetivos, realizamos uma tarefa de rememoração de sentenças (SRT) com uma tarefa de memória intermediária no formato de letras e imagens que os participantes tiveram que julgar, e a proficiência deles foi medida por meio do Vocabulary Levels Test (VLT). Coletamos dados de vinte participantes de maior proficiência, vinte e um participantes de menor proficiência e doze participantes nativos. Os dados da nossa tarefa sugerem que a proficiência foi um fator determinante para prever a habilidade dos aprendizes de L2 de corrigir sentenças agramaticais com concordância de longa e de curta distância, pois o grupo de maior proficiência superou o desempenho do grupo de menor proficiência em todas as condições alvo. No entanto, surpreendentemente, o grupo de nativos não apresentou desempenho significativamente diferente do grupo de menor proficiência em uma de nossas condições de longa distância e em uma de nossas condições de curta distância. Houve apenas uma diferença marginal entre o desempenho do grupo de maior proficiência e de menor proficiência na reprodução de sentenças gramaticais, mas detectamos uma diferença significativa entre o desempenho do grupo nativo e dos grupos bilíngues. Também encontramos diferenças significativas entre os grupos no julgamento de imagens e letras, pois os nativos foram estatisticamente melhores que os grupos bilíngues, e o grupo de maior proficiência foi significativamente diferente do grupo de menor proficiência. Discutimos nossos achados com a Hipótese do Gargalo (SLABAKOVA, 2013, 2014), o Desenvolvimento Lexical na L2 (JIANG, 2000), a Morfologia Relacional (RM) (JACKENDOFF & AUDRING, 2016), e o Modelo Fonológico/Executivo (Modelo P/E) (WEN; MOTA; MCNEILL, 2013).

PALAVRAS-CHAVE: L2; Morfologia Funcional; Memória de Trabalho; Tarefa de Rememoração de Sentenças.

LIST OF FIGURES

LIST OF GRAPHS

LIST OF TABLES

LIST OF ABBREVIATIONS

AAVE: African-American Vernacular English

BPE: Brazilian Portuguese-English

CEFR: Common European Framework of Reference for Languages

COCA: Corpus of Contemporary American English

ESL: English as Second Language

EWM: executive working memory

FIG: Figure

KET: Key English Test

L1: native language or first language

L2: second language

LMM: Linear Mixed Models

LTM: long-term memory

msec: milliseconds

NA: not applicable

NP: Noun Phrase

P/E Model: Phonological/Executive Model

PSTM: phonological short-term memory

RM: Relational Morphology

RT: reaction time

SLA: Second Language Acquisition

SRT: Sentence Recall Task

SST: speaking span test

SV: subject-verb

UCM: Unified Competition Model

VLT: Vocabulary Levels Test

VP: Verb Phrase

V2: verb-second

WM: working memory

WMC: working memory capacity

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Preliminaries

Functional morphology varies inside and between languages. Some languages have heavily inflected morphological forms, for instance, Portuguese. We would expect positive transfer from learners of languages with such a large inventory of morphological forms; however, it does not seem to be the case. While standard varieties of English use the 3rd person singular morpheme (-s) that happens in the present tense, henceforth 3rd person singular morpheme, it is not largely used by speakers of African-American Vernacular English (AAVE) (LABOV et al. since 1968). Inflectional morphology is a learning challenge across second language (L2) learners from different native languages (L1), and it has been shown to be also a processing challenge in both L2 perception and L2 production (see, for instance, Jiang, 2004, 2007, for Chinese-English learners, Mota & Baltazar (2015) and Oliveira; Fontoura & Souza, 2020 for Brazilian Portuguese-English, and Jensen et al., 2019 for Norwegian-English learners). Such challenging nature has led Slabakova (2013, 2014) to claim that functional morphology is "the bottleneck of language acquisition". According to Slabakova's hypothesis, functional morphology is one of the hardest aspects of L2 acquisition.

Functional morphology can be even more complicated if we deal with the 3rd person singular morpheme (-s) in different types of noun phrases (NP). Whenever the NP contains only the nucleus, as in local-distance agreement, it is easier for learners to make the right type of agreement, for instance, in NPs with just names, as exemplified in sentence (1) below, and pronouns, as we can observe in sentence (2). On the other hand, it is more complicated for learners to use the 3rd person singular morpheme in long-distance agreement NPs, such as in sentences (3) and (4), because learners have to isolate the nucleus of the NP. Besides, long-distance agreement sentences seem to pose an additional challenge to the less proficient learners (HOSHINO; DUSSIAS; KROLL, 2010; JENSEN et al., 2019).

(1) **Sarah** always eat**s** breakfast.

(2) **He** always drink**s** orange juice.

(3) The **engineer** with the tools always build**s** big houses.

(4) The **video** for the kids always concern**s** the parents.

Despite being quite simple from a descriptive perspective, functional morphology is a troublesome linguistic form to be mastered. Many factors can explain the underlying difficulties in functional morphology acquisition. The lack of linguistic cues in functional morphology may be a predictor of its difficulty, such as low salience. The reduced number of phones and sonority can help predict why it passes by learners undetected. Moreover, when we analyze the 3rd person singular morpheme, it conveys some redundant information that can be retrieved by other elements in the sentence (GOLDSCHNEIDER & DEKEYSER, 2001; N. ELLIS, 2017). Cognitive aspects can also offer some clarification. Working memory capacity (WMC) is very important during L2 development and production (WEN; MOTA; MCNEILL, 2013). WMC may predict L2 learners' performance in listening comprehension (AZEVEDO, 2012) and the retention and acquisition of a complex syntactic form (FINARDI, 2009).

The present study aims at contributing to the unraveling of the puzzle over the nature of this problematic issue in local-distance and long-distance agreement for L2 learners. Our contribution is based on observations of English L2 users in a written task that demands that they recall, detect, and correct violations with the 3rd person singular morpheme. We intend to shed some light on the role of WMC in the 3rd person singular morpheme perception and production. In order to pursue our study, we adapted a written sentence recall task (SOUZA, 2014; ERLAM, 2006; POTTER & LOMBARDI, 1990, 1998; LOMBARDI & POTTER, 1992; HAMAYAN; SAEGERT & LARUDEE, 1977; SACHS, 1967) to analyze if participants can recall, detect, and correct sentences without 3rd person singular morpheme. We also included an intermediary working memory (WM) task between the sentence presentation and the sentence recall to evaluate participants' WMC. This task is in the format of one-back with letters and images (KANE et al., 2007; FONTOURA, 2018). Participants had to judge whether the sequences of letters and images matched or mismatched. Since morpheme perception and production are mediated by proficiency level (HOSHINO; DUSSIAS; KROLL, 2010), we administered the Vocabulary Levels Test (VLT) to measure participants' proficiency levels. We expected that judging images and letters could be straightforward and independent from the proficiency level, but detecting ungrammatical sentences and producing the 3rd person singular morpheme would be mediated by the proficiency level.

In this study, we follow Jackendoff & Audring (2016) and their Relational Morphology (RM) theory and treat the schemas for rules generation as lexical items. Treating schemas as lexical items enables us to account for a shared network of processing and representation in long-term memory (LTM). RM claims that items that can be generated can also be stored; thus, explicit information is not deleted once automatized. Besides, each lexical item contains lemma and lexeme information inside it (LEVELT, 1989). From the perspective of an L2, the morphological information of a lexical item can only be produced if the stages of lexical development have been fully completed (JIANG, 2000). Participants may rely on resonance between new and existing patterns to learn new items (MACWHINNEY, 2017). Since we are assuming a lexical approach, our proficiency metric is also based on lexical items, meaning lemmas (NATION, 1990; LAUFER & NATION, 2001; SOUZA; DUARTE & BERG, 2015).

1.2 Statement of Purpose

The focus of our study is on functional morphology violation detection and correction. We also examine if proficiency and working memory capacity enable participants to recall, detect, and correct these violations. The 3rd person singular morpheme (-s) is investigated as it seems a special challenge for Brazilian Portuguese-English (BPE) learners (OLIVEIRA; FONTOURA & SOUZA, 2020). This type of morpheme processing and production can have a different outcome if we consider long-distance and local-distance agreement sentences, particularly for different proficiency levels (HOSHINO; DUSSIAS & KROLL, 2010; JENSEN et al., 2019). Therefore, we set four target conditions to explore the differences in long-distance and local-distance agreement sentences. The goal of our long-distance agreement conditions is to test if participants present any significant differences in detecting and correcting ungrammatical sentences with animate nouns, as it is observable in the sentence (5) below, and inanimate nouns, as in the sentence (6). Thus, we investigate whether animacy has a role in identifying and correcting these violations. On the other hand, our local-distance agreement conditions examine whether there are any significant differences in detecting and correcting ungrammatical sentences with names, as in the example in (7), and pronouns, as in (8). We contrast our ungrammatical conditions with a grammatical condition (control), as in (9), to check whether participants can retain and reproduce the 3rd person singular morpheme:

(5) *The **engineer** with the tools always ***build*** big houses.

(6) *<u>The **video** for the kids</u> always ***concern*** the parents.

(7) ***Sarah*** always ***eat*** breakfast.

(8) ***He*** always ***drink*** orange juice.

(9) Austin always **speaks** French.

Since proficiency can mediate the learners' performance, we divided our participants into lower proficiency, higher proficiency, and native groups to analyze whether proficiency is a significant factor in morpheme detection and correction in ungrammatical sentences and morpheme reproduction in grammatical sentences.

Executive functions can also have a role in Second Language Acquisition (SLA); thus, we decided to investigate if WMC could impact the three groups' performance, as WMC is used in information storage, manipulation, and comprehension (BADDELEY, 1992b; COWAN, 1999). WMC also manages learners' attention to linguistic cues in the learning process. Thus, it is essential in linguistic tasks (WEN; MOTA, MCNEILL, 2013). WMC can predict L2 learners' listening comprehension performance (AZAVEDO, 2012) and the retention and acquisition of a complex syntactic form (FINARDI, 2009). However, it may have no impact on L2 learners' morphological processing abilities (MOTA & BALTAZAR, 2015; FONTOURA, 2018). As a result, we examine if WMC has any influence on the three groups' ability to recall, detect and correct ungrammatical sentences and notice and reproduce grammatical sentences.

We implemented a sentence recall task (SOUZA, 2014; ERLAM, 2006; POTTER & LOMBARDI, 1990, 1998; LOMBARDI & POTTER, 1992; HAMAYAN; SAEGERT & LARUDEE, 1977; SACHS, 1967) with an intermediary memory task. Our task demanded that participants memorize sentences, judge sequences of images, and recall the sentence. While the grammatical sentences needed to be reproduced, ungrammatical sentences had to be corrected. The intermediary memory task was an adapted version of a one-back task (KANE et al., 2007; FONTOURA, 2018), in which participants had to judge if sequences of images and letters correspond.

We gathered data from twenty-one lower proficiency participants, twenty higher proficiency participants, and twelve native speakers. Linear Mixed Models (LMM) with binomial distributions (GODOY, 2019) were used to analyze participants' answers in the target and control sentence conditions. We divided their answers to the sentences into expected when correct and unexpected when incorrect. Our goal was to check whether the type of sentence condition influenced the participants' performance (expected/ unexpected answers) within each group. We also examined if the results between groups' performance (expected/ unexpected answers) inside each sentence condition related to their proficiency level. In addition to the previously described analyses, we ran post-hoc analyses to inspect where the effects were located within the groups and between the groups. After doing so, we analyzed if proficiency and condition type interacted. We used the LMM method to compare the sequences of images and letters that participants judged. Correct answers were considered expected and incorrect answers unexpected. The results between the groups were contrasted for this part.

1.3 Research Objectives, Questions and Hypotheses

We aim to delve into the ability that L2 learners may present in placing the 3rd person singular morpheme (-s) in a sentence recall task that demands both proficiency and WM resources in producing sentences without violations. For that purpose, we divided participants into three groups – lower proficiency, higher proficiency, and natives (control). We have four target ungrammatical sentence conditions that need to be corrected, and a grammatical sentence condition (control) that has to be reproduced. We also examine the participants' WMC through an adapted judgment of sequences of images (for more information on the one-back task, check Kane et al. (2007) and Fontoura (2018).

We conceived these specific objectives for this study:

a. Check whether there is any distinction in the performance of higher proficiency, lower proficiency, and native participants in detecting and correcting the four target ungrammatical conditions without the 3rd person singular morpheme, two with the long-distance agreement and two with the local-distance agreement, and reproducing the control grammatical condition with the 3rd person singular morpheme.

While participants need to store and reproduce the 3rd person singular morpheme in the grammatical condition, the target ungrammatical conditions demand that the information be recycled; otherwise, participants will simply parrot ungrammatical sentences without the 3rd person singular morpheme (ERLAM, 2006). Furthermore, morpheme usage is one of the hardest and last steps in L2 acquisition (SLABAKOVA, 2014). There is evidence that agreement in long-distance sentences is more complicated than in local-distance sentences considering the proficiency level (HOSHINO; DUSSIAS & KROLL, 2010; JENSEN et al., 2019). Besides, L2 learners only have access to a lexical item morphology at the end of its development (JIANG, 2000). This kind of information is integrated into a shared network of processing and representation. WM feeds on morphological information stored in long-term memory to assemble these linguistic cues in items (JACKENDOFF & AUDRING, 2016).

b. Investigate if the performance of the three proficiency groups is similar in image recall and judgment.

Sentence recall and image judgment may access different components from WM. While the first is a linguistic task, the second demands non-linguistic resources that may not be linked to proficiency. WM can be seen as an important asset in SLA (WEN; MOTA; MCNEILL, 2013). Reproducing grammatical sentences is different from correcting ungrammatical sentences with missing inflectional morphemes, especially when proficiency is at play (ERLAM, 2006). Memorizing and judging images for the three groups can be a problem (FINARDI, 2009; AZEVEDO, 2012) or not (MOTA & BALTAZAR, 2015; FONTOURA, 2018) depending on the linguistic and non-linguistic task under examination.

Considering these objectives, we envisioned the following research questions:

a. Is there any distinction in the performance of higher proficiency, lower proficiency, and native participants when contrasting their performance in reproducing a grammatical control condition with sentences inserted in a context where the 3^{rd} person singular morpheme (-s) is expected?

b. Is there any difference in the performance of the three groups when we compare their performance in detecting and correcting four target ungrammatical conditions with the long-distance agreement – animate and inanimate – and local-distance agreement – name and pronoun – without the 3^{rd} person singular morpheme (-s)?

c. Does WMC influence the groups' performance in reproducing grammatical sentences with 3^{rd} person singular morpheme (-s)?

d. Does proficiency level result in better detection and correction of ungrammatical sentences without 3rd person singular morpheme (-s)?

e. Do the three groups have a similar performance in image and judgment considering proficiency and working memory capacity?

For these research questions, we propose the following research hypotheses:

a. Grammatical sentences with 3rd person singular morpheme are more accurately reproduced than correcting ungrammatical sentences without the 3rd person singular morpheme within the groups.

According to Erlam (2006), reproducing grammatical sentences is easier than correcting ungrammatical sentences in sentence recall tasks. Reconstructing the content of the sentences and correcting violations demands more resources from participants. Moreover, Slabakova (2013, 2014) advocates that functional morphology is the bottleneck of language acquisition. Thus, it is more challenging to detect the 3rd person singular morpheme omission and place it rather than reproducing sentences with its presence.

b. Long-distance agreement ungrammatical sentences without the 3rd person singular morpheme are harder to detect and correct than the local-distance agreement ungrammatical sentences without the 3rd person singular morpheme within the groups.

Hoshino, Dussias & Kroll (2010) argue that conceptual agreement is hard to achieve in the L1 and is more troublesome in the L2. However, higher proficient L2 learners seem to overcome this difficulty. Jensen et al. (2019) defend that long-distance agreement sentences are more problematic to identify than local-distance agreement sentences. Thus, we expect that it is more complicated to spot and correct long-distance agreement ungrammatical sentences without the 3rd person singular morpheme than local-distance agreement ungrammatical sentences.

c. There is greater demand for working memory capacity for the lower proficiency group, affecting the reproduction of grammatical sentences with the 3rd person singular morpheme compared to the other groups.

Finardi (2009) and Azevedo (2012) found evidence suggesting that working memory capacity was correlated to the performance of L2 learners in the target language. Besides, Wen, Mota & McNeill (2013) claim that WMC is essential for the development and performance of the L2. We have reasons to believe that our lower proficiency group has less control of L2; therefore, WMC is a key factor for them when they recall the sentences presented to them.

d. The proficiency level influences the participants' ability to detect and correct ungrammatical sentences without the 3rd person singular morpheme between groups.

Erlam (2006) claims that higher proficiency participants showed more promising results in detecting and correcting ungrammatical sentences in sentence recall tasks.. Higher proficiency participants also displayed better performance in the studies of Hoshino, Dussias & Kroll (2010) with the conceptual agreement and Jensen et al. (2019) with long-distance and local-distance agreement sentences. Furthermore, it is only at the end of the stages of lexical development that L2 learners have access to morphological cues (JIANG, 2000). Thus, morphological information is stored in a shared framework of representation and processing. Hence, different proficiency groups have different outcomes when facing the challenge of detecting and correcting ungrammatical sentences without the 3rd person singular morpheme.

e. Proficiency does not impact image recall and judgment between the groups.

Mota & Baltazar (2015) and Fontoura (2018) discovered no relationship between the performance of L2 learners and WMC. Both studies analyzed the morphological processing abilities of L2 learners, and no significant WMC effect was found for different proficiency groups. As a result, proficiency does not affect the WMC of the groups.

1.4 Organization of the Dissertation

This first chapter introduces the subject of the study, as well as the research objectives, questions, and hypotheses that motivate it. We present the theoretical background that helps us to pursue this research in the second chapter. Chapter three regards the materials and methods upon which we are implementing this study. Results are exposed in chapter four. Finally, we discuss our findings in chapter five and the final considerations in chapter six.

CHAPTER 2

THEORETICAL BACKGROUND

This chapter is divided into the theories that support this dissertation and the studies that support the methods used. The theories are presented first, followed by the studies on the methods.

2.1 Theories

In this part, we review relevant literature to the present study. We start by talking about WM and its links to LTM concerning two models (COWAN, 1988, 1999; BADDELEY, 2003), and we also discuss a model that aims to integrate SLA and WM (WEN; MOTA, MCNEILL, 2013). Analyzing the role of LTM will open the door to talk about explicit and implicit language knowledge. When we analyze L2 knowledge, it can start as explicit or implicit language knowledge; however, it can also start as explicit language knowledge and be turned into implicit language knowledge; therefore, we will talk about explicit and implicit language knowledge according to different views (KRASHEN, [1981] 2002, [1982] 2009; DEKEYSER, 1998; R. ELLIS, 1994).

Furthermore, we review two studies concerning language learning models that approach the role of explicit language knowledge in L2 learning (KRASHEN, [1981] 2002; BIALYSTOK & FROHLICH, 1977). After it, we discuss the acquisition of L2 inflectional morphology considering an acquisition order (KRASHEN, 1977) and the apparent underlying difficulties (GOLDSCHNEIDER & DEKEYSER, 2001; N. ELLIS, 2017). As a consequence, we examine the Bottleneck Hypothesis (SLABAKOVA, 2013, 2014) and analyze the previous studies that reported the challenge that L2 learners face in morpheme processing and production (JIANG, 2004, 2007; CARNEIRO, 2008, 2011; HOSHINO; DUSSIAS; KROLL, 2010; MOTA & BALTAZAR, 2015; MACWHINNEY, 2017; JENSEN et al., 2019; OLIVEIRA; FONTOURA & SOUZA, 2020). A study found no relation between WMC and English regular and irregular past verbs processing (MOTA & BALTAZAR, 2015), but others found a correlation between WMC and retention and acquisition of a complex syntactic structure (FINARDI, 2009) and between WMC and listening comprehension performance (AZEVEDO, 2012).

We describe how the Relational Morphology (JACKENDOFF & AUDRING, 2016) regards the different mental faculties being integrated and the representation and processing as part of a shared network. In this theory, schemas for placing morphemes are part of declarative knowledge and have the same status as words. Therefore, items that can be generated can also be stored. Moreover, lexical entries contain semantic and syntactic information in its lemma and morphological and phonological information in its lexeme (LEVELT, 1989). Nevertheless, the L2 acquisition follows a different path (JIANG, 2000) because of the links between L1-L2. These shared links help the learning process (MACWHINNEY, 2017). As a result, producing morphemes in the L2 can only happen provided that L2 lexical development has been successful.

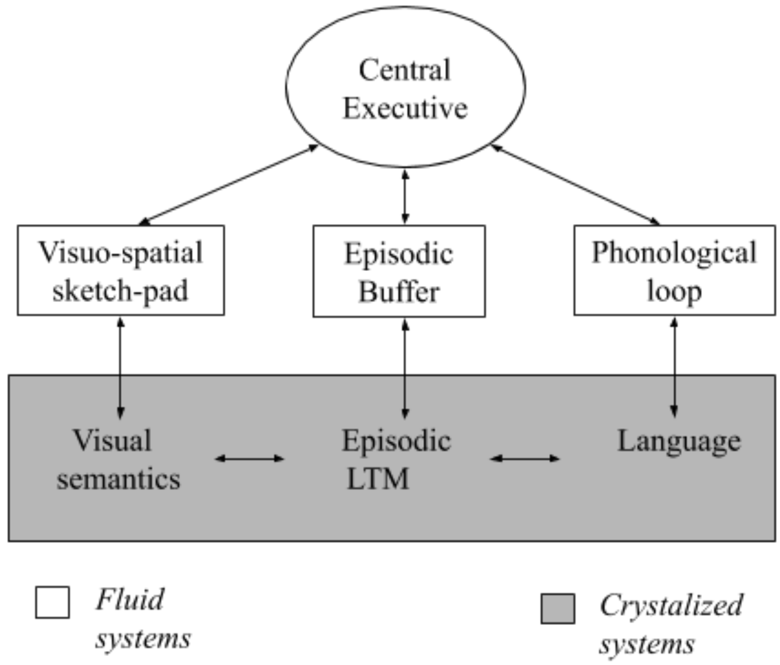
2.1.1 Working Memory Models

WM concerns information manipulation and storage; it keeps information at an accessible state so that we can perform a vast repertoire of tasks. It enables us to perform cognitive tasks such as learning, comprehension, and reasoning (BADDELEY, 1992b; COWAN, 1999).

There has been a consensus over our WM capacity limitation and its temporary storage limitation (see, for instance, MILLER, 1956; BADDELEY & HITCH, 1974; BADDELEY, 1992, 2003; COWAN, 1988, 1999, 2010). Nonetheless, there is a large variance to the extent of this capacity, 7±2 (MILLER, 1956) and 3 to 5 (COWAN, 2010), for example. We cannot forget to account for the individual differences that we may find between subjects.

Furthermore, there are many WM models from different researchers. While Baddeley & Hitch's (1974) and Cowan's (1999) models are the most well-known, they have some practical differences. Firstly, Baddeley & Hitch's (Op. cit.) model does not account for the link between WM and LTM, an issue that was corrected in Baddeley (2003), in which there seems to be an interaction between long-term systems (crystallized systems) and WM systems (fluid systems). Besides, Baddeley's (Op. cit.) model (FIG 2.1) is a multi-component model that regards processing as modular but incremental. On the other hand, Cowan's (1988, 1999) embedded-processes model of WM is a unitary model that establishes links between memory and attention.

Figure 2.1: The relationship between WM and LTM adapted from Baddeley (2003, p. 203)



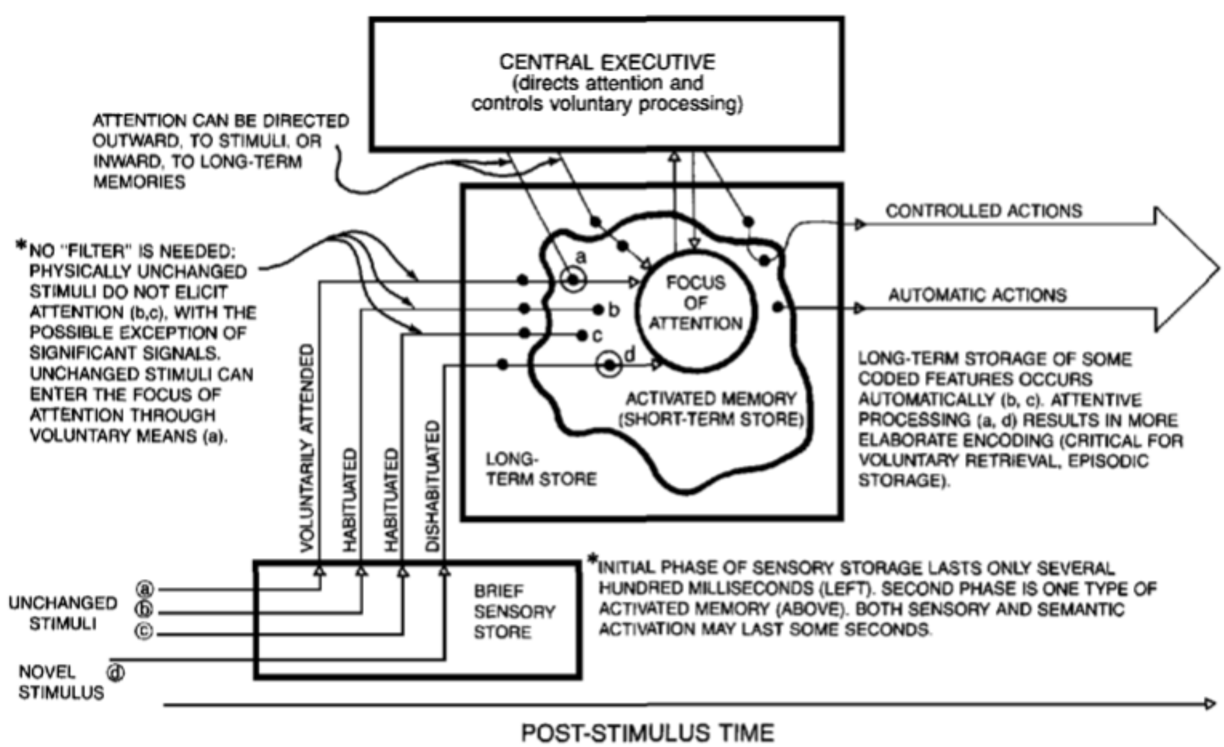
The nature of information handled by WM and LTM is different. While information on WM can be rehearsed, and therefore, kept in an active state, information on LTM has to be retrieved through associations. It means that short-term storage can rely on phonetic properties, such as rote rehearsal, and long-term storage on semantic properties, such as memory elaboration. WMC may be constrained by properties of activation, which dictate the amount of information that can be activated and kept in the focus of attention at once (COWAN, 1988).

Short-term storage refers to temporarily activated items that can be retrieved whenever attentional processes come to awareness. Information processing has to follow a complex process: (i) when faced with a stimulus, the cognitive system accesses information from LTM; (ii) the activated information gets encoded into short-term storage; (iii) new information is stored into LTM either automatically, or with the support of attention; and (iv) finally, information gradually dissolves (COWAN, Op. cit.).

Attention is capacity-limited, while activation is time-limited. “The focus of attention is controlled by voluntary processes (a central executive) and involuntary processes (the attention orienting system)” (COWAN, 1999, p. 62). Awareness impacts processing: it boosts the encoded features in perception and enables new episodic representations for explicit recall in memory. In this model, WM is supported hierarchically by: (i) LTM; (ii) the subset of activated LTM at the moment; and (iii) the subset of activated memory limited by the focus of attention and awareness. Even though this model regards LTM, it assumes that stimuli – with no importance to the activated items in memory – cannot be explicitly elicited.

In Cowan's (1988) model (FIG 2.2), information processing is embedded rather than sequential, which means that the items are intricate to each other instead of being modular. LTM information is encompassed in the large rectangle, while the subset of activated memory is encompassed in the uneven shape. Information held in the focus of attention or conscious awareness is in the small circle in the middle. According to Cowan (1999), the latter cannot be in core if it is not activated. Besides, this information can turn into new entries in LTM, such as new words or episodes.

Figure 2.2: Cowan's (1988, p. 180) embedded-processes model

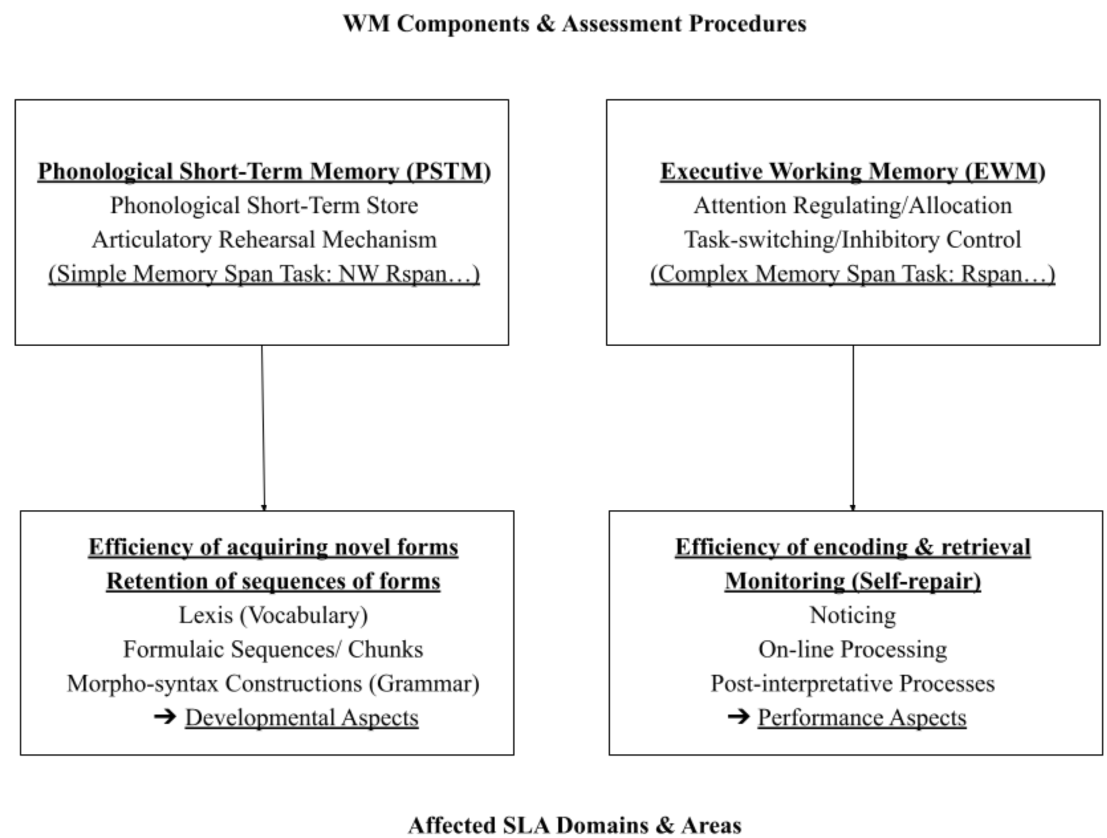


Despite Baddely's (2003) and Cowan's (1999) models being the most widely studied, Wen, Mota & McNeill's (2013) model is very promising for those studying WM and SLA because they propose a theory to unify WM and SLA. Their model aims to characterize and measure WM resources that are specifically used in SLA domains and tasks. Their Phonological/Executive Model (P/E Model) is based on Ullman's (2012) Declarative/Procedural long-term memory model and is an attempt to solve the dichotomy between the focus on the phonological short-term memory (PSTM) (see, for instance, Baddeley & Hitch (1974); Baddeley (1992, 2003)) and the executive working memory (EWM) (check Cowan (1988, 1999, 2010)). There are two levels in the P/E Model: (i) WM components, including mechanisms and functions, and assessment procedures related to them; and (ii) SLA domains and areas that can be affected by the WM components.

The first level encompasses the mechanisms involved in both PSTM and EWM (FIG 2.3). While PSTM deals with phonological short-term storage and articulatory rehearsal mechanisms, EWM is concerned with attention regulating and allocation associated with task-switching and inhibitory control. The second level describes SLA domains and areas that can be affected by the WM components in the first level. PSTM results in acquiring new word forms and recalling sequences of forms, such as vocabulary, chunks, and morpho-syntactic constructions. It is an asset linked to the developmental stages of SLA. The outcome of EMW is encoding and retrieval of language forms in on-line language processing and post-interpretive processes. Therefore, one needs EWM for monitoring and self-repair, which are critical for speech and written performance, and cognitive functions such as noticing and post-interpretative processes.

According to this model, three types of WM effects can be investigated: main effects of PSTM or EWM, the interaction effects of PSTM or EWM with other elements, and/or threshold effects, such as WMC levels that enable L2 learning to happen smoothly. Besides WM effects, proficiency in the L2 should also be considered. WM tasks should be chosen based on the participant’s proficiency level; thus, more complicated tasks should be assigned only to high proficient participants.

Figure 2.3: The Phonological/Executive Model (WEN; MOTA; MCNEILL, 2013, p. 7)



Our task demands that participants recall, encode, and retrieve linguistic information because they have to recall grammatical sentences and recall and correct ungrammatical sentences. Participants need specific WM resources related to SLA to process, store, manipulate, and correct the linguistic stimuli. Besides, we are dealing with L2 learners' performance aspects that involve rehearsal mechanisms, noticing, attention regulation, and task-switching. In addition to the sentence recall, we have an intermediary memory task that participants have to judge sequences of images (KANE et. al., 2007; FONTOURA, 2018). Since we are dealing with non-linguistic and linguistics WM resources, we believe that Wen, Mota & McNeill's model can help explain the outcomes produced by participants in the sentence recall part and in the judgment of images.

2.1.2 Explicit and Implicit L2 Knowledge

It appears that there is a debate in the literature over the nature of language knowledge and whether non-native learners can turn their explicit knowledge into implicit knowledge[1]. We can think about the relationship between explicit and implicit knowledge considering three approaches: the non-interface position, the strong interface position, and the weak interface position.

The non-interface position by Krashen ([1981] 2002, [1982] 2009) assumes that explicit and implicit language knowledge are independent of one another, as "learning does not "turn into" acquisition" (KRASHEN, [1982], 2009, p. 83). The author argues that L2 learners, who know explicitly many rules, still make "careless" mistakes in production. This occurs because the learner knows the rule but has not internalized it. Many late L2 learners may know the rules of 3rd person singular verbs – e.g. *He goes to work every day* (KRASHEN, [1982], 2009, p. 83) – but cannot produce it. Thus, explicit knowledge of rules does not account for fluency.

The strong interface position assumes that "explicit knowledge can be automatized through and for production" (DEKEYSER, 1998, p. 58). Learners go through stages of acquisition, in which they start with the support of declarative knowledge up to its proceduralization and refinement. Learners do not exchange their declarative knowledge for procedural knowledge but depart from declarative exclusively towards procedural knowledge. Recurring behaviors, such as rules practice, may result in the restructuring of declarative knowledge and, consequently, in proceduralization. Elements that co-occur can be turned into larger chunks reducing the WM load and making the processes easier and faster.

According to the weak interface position, "explicit knowledge derived from formal instruction may convert into implicit knowledge, but only if the learner has reached a level of development that enables her to accommodate the new linguistic material" (R. ELLIS, 1994, p. 88). This means that the learner's explicit knowledge can be incorporated into the interlanguage in the form of implicit knowledge if explicit knowledge is ready to move to another step in the developmental process. In this position, formal instruction is seen as a way to help learners enhance their control in the L2, and therefore, as a means to have explicit knowledge turned into implicit knowledge. Furthermore, for this position, most of the time, knowledge does not begin as explicit

---

1 In this study, we consider that implicit knowledge and procedural knowledge are equivalent, and therefore, explicit knowledge and declarative knowledge are also the same

but rather as implicit knowledge.

Considering the three interface positions, we follow the strong interface position as it assumes that items can be stored in the declarative memory along with the procedural memory. Thus, items are not deleted from declarative memory once automatized, and they can be fully specified in memory. Whenever items reach a high level of resting activation, they will reduce the WM load and will be activated more easily (DEKEYSER, 1998). WM is also a meaningful asset in SLA and performance processes (WEN; MOTA; MCNEILL, 2013). In this study, we investigate if L2 learners can recall, detect, and correct violations with the inflectional morpheme -s, 3rd person singular, in a sentence recall task. Therefore, it is essential to consider if the learners can use their implicit and explicit knowledge to correct written sentences. In addition, WM components have an important role in storing and manipulating linguistic stimuli for L2 learners. Should an item have a higher level of activation, L2 learners can retrieve it more easily. We follow this strong interface position about explicit and implicit language knowledge because it seems to be in accordance with the other theories we chose as our theoretical background.

2.1.3 L2 Learning Models

According to the Acquisition-Learning Distinction[2] by Krashen ([1982] 2009), there are two different and independent ways in which adults develop competence in an L2. The first concerns language acquisition, a process that resembles pretty much how children acquire their L1. That is, language acquisition is subconscious in a way that learners are not usually aware of the underlying rules of languages. The author also calls this type of acquisition “implicit learning, informal learning, and natural learning” (p. 10). The second way of developing competence in an L2 refers to language learning, also called explicit learning, which is obtaining a language consciously; in other words, being able to talk about the rules of the language.

The process of language acquisition is analogous in both first and second languages, in the sense that it needs interaction in both languages. In these situations, speakers are mostly worried about the message and not about the formal aspects. The Monitor Hypothesis by Krashen ([1981] 2002) predicts that conscious learning can be used by learners just as a Monitor, whereas the

[2] Krashen’s ([1982] 2009) theory of SLA distinguishes among five hypotheses: the acquisition-learning distinction, the natural order hypothesis, the monitor hypothesis, the input hypothesis, and the affective filter hypothesis. In this study, we discuss some of the assumptions related to the first three hypotheses, namely, the acquisition-learning distinction, the natural order hypothesis, and the monitor hypothesis.

acquired system is responsible for active communication. Thus, explicit knowledge of the language can be used as a device to modify the output before or after it is spoken. When it comes to L2 production by adults, the author presents a model in which the acquired competence and the learned competence are interrelated in adult L2 production, as we can check in FIG. 2.4:

Figure 2.4: Model for Acquisition and Learning in L2 Adult Production by Krashen ([1982] 2009, p. 16):

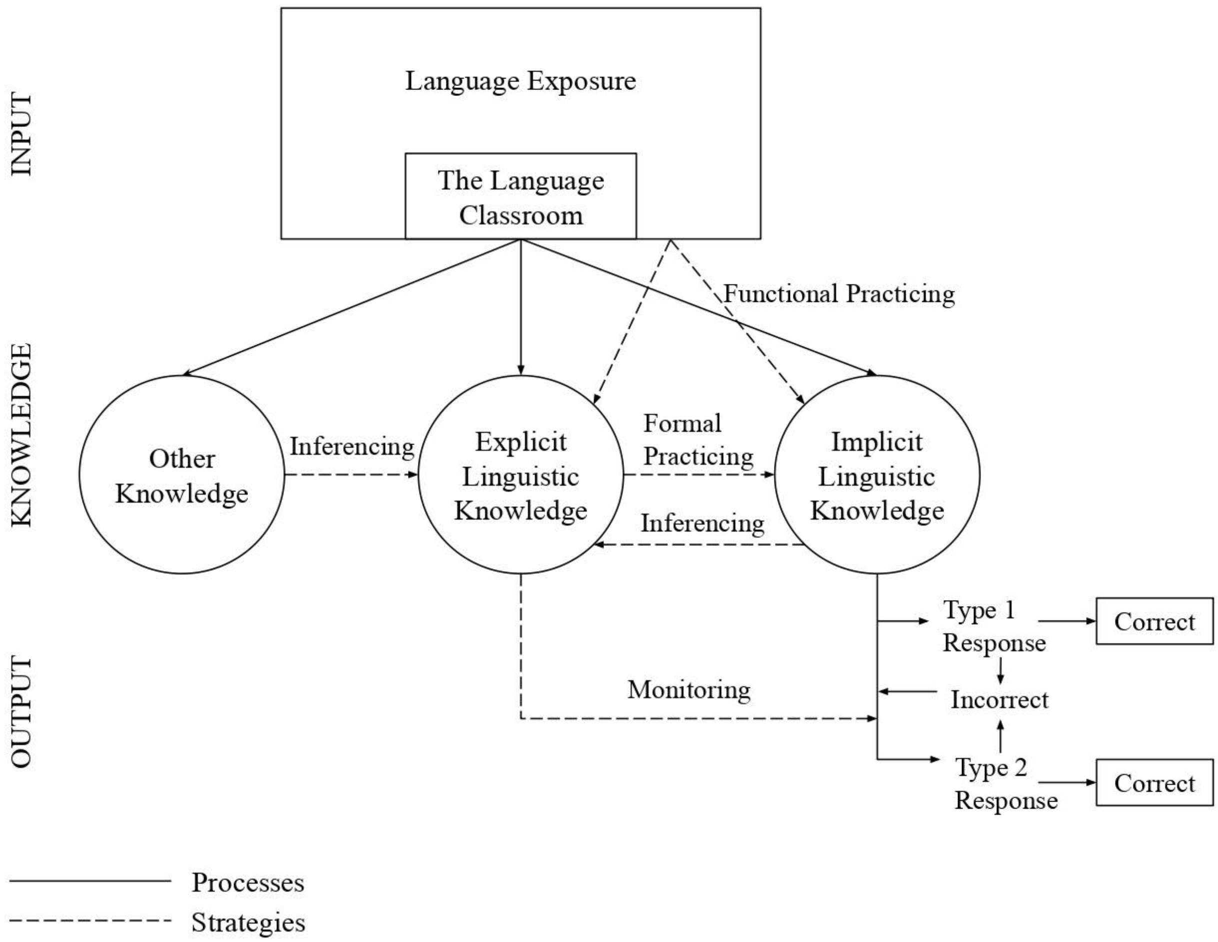
In this hypothesis, formal rules, namely conscious learning, have a restricted function in L2 performance. The Monitor has some constraining conditions to be used. The first condition deals with time; the performer can use the Monitor, as long as there is time to employ it. During interactions, this can be especially difficult. Thus, the Monitor has a slight or no effect under such circumstances. The second condition regards the performer's ability to focus on form or accuracy, which is difficult to attend. The third condition to using the Monitor is the necessity to know the rule; without having a mental representation of it, one cannot apply it. The author defends that it is challenging to use conscious learning as the three conditions are hardly ever satisfied.

Bialystok & Frohlich (1977) also aimed to develop a model for second language learning, as illustrated in FIG. 2.5 below. Their model has three levels of organization – Input, Knowledge, and Output. Input encompasses a general framework called Language Exposure, from which it is possible to detach certain language experiences, such as the Language Classroom, books, trips, etc. Language Classroom conveys information to all three knowledge sources. In this environment, learners have contact with both formal and functional aspects of language. The Knowledge level involves three different types: Other Knowledge, Explicit Linguistic Knowledge, and Implicit Linguistic Knowledge. The first type accounts for knowledge about the world and other languages rather than the target language; the second type relates to conscious knowledge of L2 formal aspects – e.g., syntax; and the third type of knowledge comprises unconscious proficiency in the target language. Output, the third level, includes both comprehension and production of the target language.

This model can be operationalized following three parameters: learning processes, learning strategies, and learner factors. Learning processes “are concerned both with the way in which the three knowledge sources are built up and utilized for specific language tasks and with the mechanisms underlying the production of responses” (BIALYSTOK & FROHLICH, 1977, p. 4). Learning strategies encompass practicing, monitoring, and inferencing. It refers to the relationship between the different types of knowledge sources and language outcomes[3]. Practicing happens whenever the target language is used. It is divided into formal and functional practices, which are different strategies. Formal practice is the transference from Explicit to Implicit Knowledge. Functional practice can be described as the amount of language exposure and classroom encounters, being the language outside the class as the source of functional practice. In their study, functional practice enhanced performance in formal and functional tasks. Monitoring has been adapted from the monitor model of Krashen (1976), which means that it uses explicit knowledge in order to formulate/modify the output. According to Carton (1971), inferencing is “a process of identifying unfamiliar stimuli. In foreign language learning, inferencing is concerned with the acquisition of new morphemes and vocables in “natural” contexts” (p. 45). Inferencing should not be confused with logical inference, as the former deals with “the implications of convergent cues” while formality levels are not required. In Bialystok and Frohlich’s model, inferencing gets the significant information from either Other Knowledge or Implicit Linguistic Knowledge and takes it to Explicit Knowledge, being aware to the learner. Learner’s individualities regulate how the model shall work.

[3] These parameters are conscious ways in which the learner can improve learning and enhance proficiency (BIALYSTOK & FROHLICH, 1977).

Figure 2.5: Revised Model of Second Language Learning adapted from Bialystok & Frohlich (1977, p. 66)[4]:



Both models address crucial aspects of L2 learning and account for different types of knowledge and learning. An L2 model should regard how explicit knowledge can be turned into implicit knowledge. Different learning contexts may present differences in dealing with explicit knowledge. While Krasen ([1981] 2002, [1982] 2009) seems to work with more "naturalistic" settings of language acquisition, Bialystok & Frohlich (1977) study L2 acquisition from a formal

[4] The authors depicted one model before conducting their study, but after getting some data from the study, some revisions had to be made. The model presented here already accounts for the revised model.

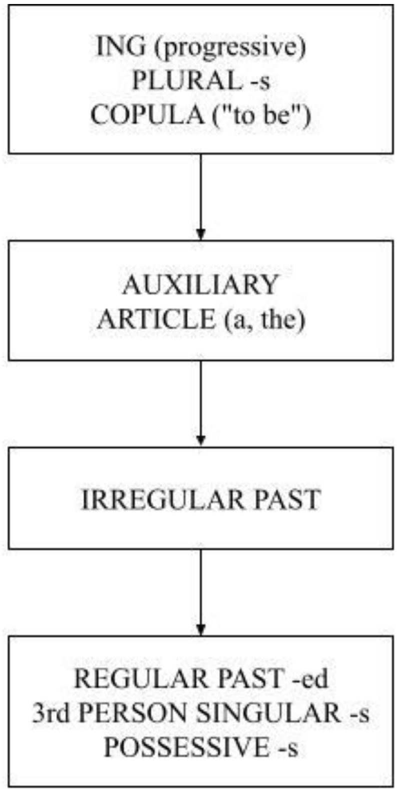
instruction view. In the present study, most of our participants received thorough training in L2 in language classrooms. In addition, our task demands that participants correct sentences with violations; they have a fair amount of time to do so (100,000 milliseconds). However, the Monitor could only be used if participants could focus on the form and knew the rule; time is just one of the requirements, as suggested by Krashen ([1982] 2009). Consequently, we aimed to test whether L2 participants could recycle information in WM and apply their explicit knowledge; when compared to L1 participants executing the same task.

2.1.4 L2 Morpheme Acquisition

In what concerns the acquisition of grammatical structures by L2 learners of English, Krashen (2009 [1982], p. 12) defends that it "proceeds in a predictable order"[5]. While some grammatical structures are easily acquired, others pose some challenges. Although there may be some variance in the order of L2 morpheme acquisition, the author defends that it happens significantly similarly. In what comes to the natural order of morpheme acquisition, Krashen (1977) proposes a model in which it is possible to observe the easiest and hardest morphemes to acquire for adult L2 learners, as observable in FIG 2.6:

[5] Dulay and Burt (1974, 1975) also claim a "natural order" of morpheme acquisition for children in English SLA.

Figure 2.6: Natural Order of L2 Morpheme Acquisition for adults, adapted from Krashen (1977, p. 149):



While the progressive form (-ING), the plural (-s), and the copula (to be) seem to be very straightforward to acquire, others such as the regular past (-ed), the 3rd person singular morpheme (-s), and the possessive (-s) are very troublesome. Krashen (1977) claims that the forms in each stage of this model are acquired at about the same time. The author proposed this order of acquisition based on a study that he conducted with others. In this latter study, they conclude that their "data, then, is quite consistent with the hypothesis that grouped adult free speech, using "intermediate" level ESL performers, does produce[6] a "natural order" for the set of grammatical morphemes" (KRASHEN; HOUCK; GIUNCHI; BODE; BIRNBAUM & STREI, 1977, p. 340).

Nonetheless, DeKeyser (1998) contests this "natural order" of acquisition, as many of these studies were performed on non-instructed learners. This could have had a different outcome if the participants had had instruction on the structures. On the other hand, studies that were done with instructed learners usually have little information about how the instruction was presented.

[6] Since we have a production task in our study, we believe it is critical to attempt to understand more about the factors that can hinder morpheme production.

Goldschneider & DeKeyser (2001) analyzed many studies that advocate a "natural order" of L2 morpheme acquisition and attempted to find the underlying cause for this apparent difficulty. They followed Brown's (1973) guidelines for L1 acquisition. For them, perceptual salience, semantic complexity, morphophonological regularity, syntactic category, and frequency can explain why this is such an arduous form to be mastered.

The twelve selected studies comprise data from L2 learners of twenty-eight diverse L1 language families; therefore, they assumed that L1 diversity was well balanced and the L1 transfer was excluded from their meta-analysis. Their multiple regression analysis indicates that the variance in L2 morpheme acquisition can be largely explained by these factors, and saliency has the strongest impact.

Earlier acquired forms are usually very perceptually salient, whereas forms that are hard to perceive are harder to learn. The authors attribute three factors to be accounted for in perceptual salience: the number of phones (phonetic substance), whether a surface form has a vowel (syllabicity), and the amount of sonority in the form. Thus, a higher amount of phones, a vowel in the surface syllable, and a more sonorous sound should yield more perceptual salience.

Semantic complexity handles the allocation of meanings expressed by a form. Take, for instance, the plural marking (-s) that deals with number and the 3rd person singular marking (-s) that deals with person, number, and present tense. Forms that display more meanings can be harder and take longer to acquire in comparison to more straightforward forms.

Morphologically regular forms are less affected by their phonological environment, and therefore, they exhibit less allomorphy[7]. Besides, factors such as contractibility can also play a role in variability. Consequently, the number of phonological alternations and homophony with other forms can be decisive to determine the scale of acquisition. The allomorphs /s, z, əz/ of the plural (-s), possessive (-'s), and 3rd person singular morpheme (-s) may be a special challenge to be acquired. Moreover, -s is homophonous with three forms, plural (-s), possessive (-'s), and 3rd person singular morpheme (-s), which is another great difficulty to be faced. Another factor that comes into play is redundancy; in other words, how much one needs to use a certain form to convey meaning. The meaning encoded in the 3rd person singular marking (-s) can be retrieved by other

---

7 This study has not investigated the difference between morphological and phonological acquisition in L2 learners. This salient difficulty could be related to the syllable configuration and phonotactic rules in regular morphemes' pronunciation.

elements in the sentence, such as by the subject and a time expression. According to N. Ellis (2017, p. 25), "L2 learners have been found to prefer adverbial over inflectional cues to tense in naturalistic SLA".

In what concerns the syntactic category, Goldschneider & DeKeyser (2001) follow Zobl and Liceras (1994), who divide the forms regarding their lexical or functional category. "For L2 acquisition, the ranking of the morphemes depends on the bound/affixal exponent versus free exponent contrast, with the free exponents and the bound exponents ranking at opposite ends of the hierarchy" (ZOBL & LICERAS, Op. cit., p. 172). While they regard 3rd person singular (-s) and regular past (-ed) morphemes as functional affixes, because of their syntactic content, plural (-s) and possessive (-'s) are considered lexical affixes since they are found along with lexical words. As a result, functional categories are decisively harder in the acquisition scale as lexical affixes, such as progressive form (-ING) and plural (-s), are not acquired late by L2 learners. As reported earlier, the possessive (-'s) can also pose a problem for L2 learners; it seems that lexical categories can contain its difficulties. Although this is an important justification for them, Goldschneider & DeKeyser (2001) did not find substantial evidence for this division.

Frequency in the input has long been advocated for its role in L2 acquisition. The amount of exposure one has to a certain structure may affect its learnability. Larsen-Freeman (1976) attempted to analyze the "fixed order of morpheme acquisition". She administered five tasks – reading, writing, listening, imitating, and speaking – to a group of twenty-four adult L2 learners. A strong correlation was found between frequency and the morphemes acquired, "the tentative conclusion is that morpheme frequency of occurrence in native-speaker speech is the principal determinant for the oral production morpheme order of second language learners" (LARSEN-FREEMAN, Op. cit., p. 132). However, not all L2 learners are immersed in the L2 context and the type of input L2 learners receive can also be at play.

Furthermore, N. Ellis (2017) delved into the role of psychophysical salience, contingency of form function-association, and learned attention in L2 acquisition. These factors would help to clarify why grammatical morphemes and closed class constructions are more troublesome to learn than open class constructions. Bottom-up auditory processing is most times insufficient to perceive grammatical functors. L1 learners can perceive these subtle cues because their systems enable them to use top-down processing, which L2 learners lack. Therefore, low salience, redundancy, low contingency in form-function mapping, and adults' L1 knowledge biasing L2 learning can help explain this difficulty.

Nonetheless, they also found that there are other factors that should also be regarded, such as perceptual salience, semantic complexity, morphophonological regularity, syntactic category, and frequency for Goldschneider & DeKeyser (2001) and psychophysical salience, contingency of form function-association and, learned attention in L2 acquisition for N. Ellis (2017).

Additionally, Mota & Baltazar (2015) conducted a study to analyze the effects that frequency, proficiency, WM, and inhibitory control can exert in English regular and irregular past verbs processing for BPE. Participants were divided into low proficiency, high proficiency, and natives (control group). All of the groups carried out three tasks: (i) Frequency Effect Task, (ii) Simon Task with arrows, and (iii) Letters and Numbers Ordering Task. In the first task, all stimuli were visually displayed at once. A verb in the infinitive form without *to* was presented followed by a sentence in the simple present tense. Finally, participants saw another sentence in which they needed to orally produce a verb in the simple past tense, as exemplified in (10):

(10) STUDY

Every day I study English.

Yesterday I ________________ English.

(MOTA & BALTAZAR, 2015, p. 141)

Both the second and the third tasks investigated non-linguistic information. The Simon Task aimed to check participants' ability to inhibit irrelevant information. In the incongruent condition, the arrow pointing direction and the position it appeared on the screen mismatched; that is, the first would be pointing right while it appeared on the left side. The goal of the Letters and Numbers Ordering Task was to examine participants' WMC. Sequences of letters and numbers were visually and orally introduced to participants in their native language. After seeing the stimuli, participants had to orally rearrange the numbers in ascending order and then the letters in alphabetical order, as we can check in table 2.1:

Table 2.1: Letters and Numbers Ordering Task from Mota & Baltazar (2015, p. 141)

Stimuli	6–G–A–8–X
Expected Outcome	6–8–A–G– X

In what concerns the first task, the analysis within the groups presented no significant difference in their reaction time (RT) when comparing irregular verbs with high and low-frequency. However, they found a statistically significant difference within the high proficiency and low proficiency groups' RT in the comparison between high and low-frequency regular verbs. The analysis within the low proficiency and high proficiency groups displayed substantial differences in accuracy in producing both regular and irregular verbs, contrasting low and high-frequency verbs. No significant difference was found in the native's group accuracy. Thus, higher frequent regular and irregular verbs were more easily conjugated for the bilinguals.

Statistically significant results were found when contrasting the RT between the groups for regular and irregular verbs. The natives were the fastest, while the high proficiency group was faster than the low proficiency group in producing inflected regular and irregular verbs, with low and high-frequency. Proficiency was seen as an important asset in processing. The accuracy results between the groups were more divergent. No significant difference between the groups' high-frequency verbs production accuracy was found, but there was a significant difference between the native and low proficiency groups and the high proficiency and low proficiency groups in low-frequency verbs production accuracy. They also found a difference between the groups in high-frequency and low-frequency irregular verbs regarding the comparison between the natives to the low proficiency group, the high proficiency group to the low proficiency group, and the natives to the high proficiency group. Besides, the low proficiency group exhibited a significantly less accurate result for irregular low-frequency verbs than the other two groups. Proficiency was very important to determine the accuracy in producing low-frequency verbs. Nonetheless, the bilingual groups' performance on the first task was not related to their performance on the second and third tasks. Therefore, it is not an inhibitory control or WMC impairment. The results concerning the effects of regular and irregular inflected English verbs frequency processing in Mota & Baltazar (2015) were inconclusive. The frequency effects located in regular verbs indicate that the bilinguals may store them as a whole unity. Fontoura (2018) found no meaningful difference between higher proficiency and lower proficiency groups in acceptability judgment tasks with memory load and 2-back tasks. In her study, inflectional morpheme processing was not related to WMC. In Mota & Kramer (2015), early and late bilinguals outperformed monolinguals in tasks that demanded WMC and executive control. The results suggest that bilingualism presented a substantial effect despite the context and the age of acquisition.

The core of Mota and Baltazar's (2015) study was regular and irregular past tense verbs processing by BPE and natives. They reviewed a dual-route model (PINKER, 1991[8] *apud* MOTA & BALTAZAR, 2015, p. 135) to decompose inflected past verbs and store irregular verbs, but they also regarded a unitary model for language processing (SEIDENBERG; MCCLELLAND, 1989[9] *apud* MOTA & BALTAZAR, 2015, p. 135). Their results could not be explained by the dual-route model, as bilinguals did not seem to decompose regular verbs, and yet, they say there is very little evidence for a unitary model that processes both regular and irregular verbs the same way. Furthermore, another possibility of interpretation could be the study of Jackendoff & Audring (2016)[10] to explain why their participants stored these regular verbs as a whole unity instead of decomposing them. Langacker (1987) argues that the past forms are also stored. Another important observation is that the mechanisms used for production may differ from the ones used for comprehension (JIANG, 2000). One may understand lower-frequency words they come across, but they do not have the vocabulary depth to use them (SCHIMITT, 2014). L2 learners may rely on shared links between stored verb forms. Therefore, this less automatic process may end up producing mistakes (JIANG, 2000).

Finardi (2009) investigated if WMC and the retention and acquisition of the syntactic structure (*So + aux + I* and *Neither + aux + I*) in L2 were correlated. She investigated how the agreement in Portuguese occurs by repeating the main verb, as in (A) and (B):

(A) Eu falo espanhol.

(B) Eu também falo.

(FINARDI, 2009, p. 8)

---

8 PINKER, S. Rules of language. Science, v. 253, p. 530–535, 1991.

9 SEIDENBERG, M.S.; MCCLELLAND, J.L. A distributed, developmental model of word recognition and naming. Psychological Review, 96, 523-568, 1989.

10 We will discuss this study in section 2.6.

However, in English, this agreement process needs to be submitted to a different syntactic computation by using an auxiliary verb in the same time tense of the main verb, as in (C) and (D):

(C) I speak English.

(D) So do I.

(FINARDI, 2009, p. 8)

The ideas expressed by the particles *So* and *Neither* are encapsulated and positioned in the phrase front. In order to test her hypothesis, participants carried out a speaking span test (SST) in the L1 and in the L2. The author gathered data from ninety-seven adult Brazilian-English learners; fifty were considered control and forty-seven experimental. Participants retained the structure if they used the target structure correctly in a focused immediate test. Acquisition of the target structure, on the other hand, was operationalized in an unfocused delayed test. Her results indicate that WMC, concerning SST, is related to the acquisition of a complex structure in L2 speech. WMC mediates the acquisition of a syntactic rule in L2 speech by considering controlled processes in a rule-based system.

In addition, Azevedo (2012) examined (i) if individual differences in WMC of BPE low-proficiency learners could predict listening comprehension performance in a Cambridge proficiency test, Key English Test (KET), and (ii) if two months of explicit training with a focus on listening strategies would improve the learners' KET scores. She collected data from two groups of twenty-four adult BPE learners that were enrolled in preparatory classes for KET. Firstly, twenty-four participants performed a working memory battery test (BAMT-UFMG). Secondly, fourteen participants (experimental group) underwent two months (15 classes) of explicit listening training while ten participants (control group) had no training. Both of her hypotheses were confirmed because WMC predicted participants' listening comprehension performance in KET and two months of explicit training improved the experimental group performance by 14%, whereas the lack of training made the performance of the control group decrease by 3%.

As a result, WMC can have a big impact on L2 learners depending on the task being evaluated. Since we are dealing with a task that demands WMC and proficiency from our participants, we can have different outcomes considering our linguistic task, in which participants have to recall and correct ungrammatical sentences, and our non-linguistic task, in which participants have to judge sequences of images.

2.1.5 Bottleneck Hypothesis

Slabakova (2013, 2014) argues that different linguistic features can be acquired at different times. In particular, functional morphology is one of the most complicated grammatical aspects to acquire, as it is "the bottleneck of language acquisition". That is so because functional morphology postulates formal grammatical features that are challenging in both production and comprehension. Therefore, mastering an L2 implies mastering the functional morphology of the language.

While native speakers produce an enormous array of grammatical features, "L2 learners exhibit optionality or variability in their use of verbal and nominal inflection and associated lexical items" (WHITE, 2003, p. 178). Slabakova bases her Bottleneck Hypothesis on some of the work of White (Op. cit) as she defends that syntax is more straightforward to acquire than morphology. White assumes the morphology/syntax interface in interlanguage grammars. This interface has two opposing views. The first view portrays variability in inflectional morphology being a problem of grammatical representation. This variability is due to either (i) a developmental inconsistency in the interlanguage grammar that can be later acquired; or (ii) a deficit that is everlasting. The second view regards abstract morphosyntactic features as part of an early interlanguage grammar, and therefore, an underlying syntactic representation is operational. According to this approach, sometimes, the learner is not able to access morphological information even though it has been acquired. In such a case, a breakdown in the association of one part of the grammar to another may happen (HAZNEDAR & SCHWARTZ, 1997[11]; LARDIERE, 1998a[12], b[13], 2000[14]; LARDIERE & SCHWARTZ, 1997[15]; PRÉVOST & WHITE, 2000a[16], b[17]; ROBERTSON, 2000[18] *apud* WHITE,

---

[11] HAZNEDAR, B. & SCHWARTZ, B. D. Are there optional infinitives in child L2 acquisition? In E. Hughes, M. Hughes and A. Greenhill (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language.* Somerville, MA: Cascadilla Press, *Development,* p. 257–68, 1997.

[12] LARDIERE, D. Case and tense in the 'fossilized' steady state. *Second Language Research,* v. 14, p. 1–26, 1998a.

[13] LARDIERE, D. Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research,* v. 14, p. 359–75, 1998b.

[14] LARDIERE, D. Mapping features to forms in second language acquisition. In: J. Archibald (ed.), *Second language acquisition and linguistic theory.* Oxford: Blackwell. p. 102–29, 2000.

[15] LARDIERE, D. & SCHWARTZ, B. D. Feature-marking in the L2 development of deverbal compounds. *Journal of Linguistics,* v. 33, p. 327–53, 1997.

[16] PRÉVOST, P. & WHITE, L. Accounting for morphological variation in L2 acquisition: truncation or missing inflection? In M.-A. Friedemann and L. Rizzi (eds.), *The acquisition of syntax.* London: Longman, p. 202–35, 2000a.

[17] PRÉVOST, P. & WHITE, L Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research,* v. 16, p. 103–33, 2000b.

[18] ROBERTSON, D. Variability in the use of the English article system by Chinese learners of English. *Second Language Research,* v. 16, p. 135–72, 2000.

2003, p. 179).

White (2003) depicts two perspectives on the L2 morphology/syntax interface. The first proposal is the morphology-before-syntax (or Rich Agreement), which argues that the lack of morphological display is seen as a sign of missing syntactic properties from the interlanguage grammar, momentarily or permanently. The second proposal is the syntax-before-morphology (or Separation Hypothesis), which defends that overt morphological cues become part of one's knowledge after the syntax. In this view, the most crucial matter is the morphosyntactic features stored in the interlanguage grammar rather than learners presenting the right morphology (EPSTEN *et al.,* 1996[19]; HAZNEDAR & SCHWARTZ, 1997[20]; IONIN & WEXLER, 2002[21]; LARDIERE 1998a[22], b[23], 2000[24]; LARDIERE & SCHWARTZ, 1997[25]; PRÉVOST & WHITE, 2000a[26], b[27]; SCHWARTZ, 1991[28] *apud* WHITE, 2003, p. 188). Slabakova adopts the latter view and bases her theory on it.

Based on this, Slabakova (2014) states the reasoning behind the bottleneck hypothesis (FIG. 2.7): i) functional morphology reveals distinctions in syntactic and semantic properties between languages; ii) syntactic and semantic relations are universal; iii) functional morphology dictates the acquisition of syntax and semantics in the L2; iv) morphology is the bottleneck of language acquisition

---

[19] EPSTEIN, S.; FLYNN, S. & MARTOHARDJONO, G. Second language acquisition: theoretical and experimental issues in contemporary research. *Brain and Behavioral Sciences, v.* 19, p. 677–758, 1996.

[20] See 11 above.

[21] IONIN, T. & WEXLER, K. Why is 'is' easier than '-s'?: acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research,* v. 18, p. 95–136, 2002.

[22] See 12 above.

[23] See 13 above.

[24] See 14 above.

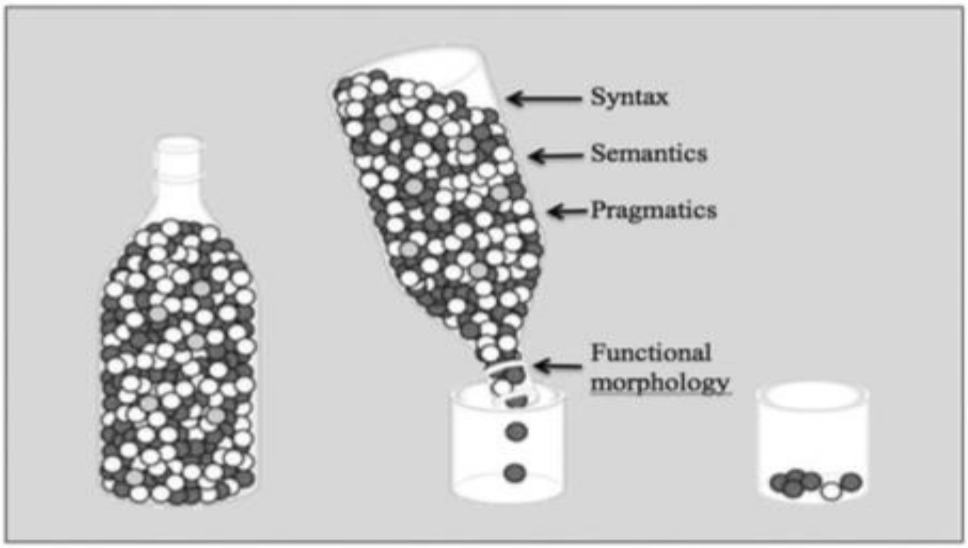
[25] See 15 above.

[26] See 10 above.

[27] See 11 above.

[28] SCHWARTZ, B. D. Conceptual and empirical evidence: a response to Meisel. In: Eubank L. (Ed.), *Point Counterpoint: Universal Grammar in the second language.* Amsterdam: John Benjamins, p. 277–304, 1991.

Figure 2.7: The Bottleneck Hypothesis of L2 Acquisition by Slabakova (2014, p. 547):



Furthermore, Jensen et al. (2019) used an untimed acceptability judgment task to compare verb-second (V2) order and subject-verb agreement in long-distance sentences with singular subjects, as illustrated in the sentence (11), and plural subjects, as in (12), and local-distance sentences with singular subjects, as it is possible to observe in the sentence (13), and plural subjects, as in (14), for Norwegian-English bilinguals. Participants saw both the grammatical and the ungrammatical versions of the same sentence. Sentences' length had 10 to 12 syllables and contained the top 5,000 words from Corpus of Contemporary American English (COCA):

(11) Long-distance singular agreement
The teacher with black shoes walks to work every day.
*The teacher with black shoes walk to work every day.

(12) Long-distance plural agreement
The kids with the red bike play in the garden.
*The kids with the red bike plays in the garden.

(13) Local-distance singular agreement
The brown dog plays with the yellow football.
*The brown dog play with the yellow football.

(14) Local-distance plural agreement

The teachers give their students a lot of homework.

*The teachers gives their students a lot of homework.

(JENSEN et al., 2019, p. 12)

While V2 ungrammaticality was identified for both less and more proficient groups, the subject agreement was problematic for both groups, in long-distance agreement sentences more than in local-distance agreement sentences. Yet, morpheme oversuppliance was harder to spot than its omission, particularly in long-distance subjects. The authors attribute the 3rd person singular morpheme (-s) low salience and redundancy as predicting factors for its difficulty. They argue that these results are tentative evidence supporting the bottleneck hypothesis.

Hoshino, Dussias & Kroll (2010) studied how English-Spanish bilinguals (English native speakers living in an English-speaking environment who spoke Spanish as an L2) and Spanish-English bilinguals (Spanish native speakers who spoke English and lived in the United States) completed sentence fragments. They aimed to explore how the participants would use grammatical and conceptual number. Participants saw sentence preambles consisting of an NP with a singular noun followed by a prepositional phrase (PP). In one item, there was a mismatch in the single-referent, as exemplified in (15 a), and, in the other, there was a match, as in (15 b). The same was true for a pair of preambles with distributive-referent that mismatched, such in the sentence in (16 a), and others that matched, as in (16 b). Only one item from the pair was displayed to each participant:

(15 a) The author of the novels

(15 b) The author of the novel

(16 a) The drawing on the posters

(16 b) The drawing on the poster

(HOSHINO; DUSSIAS; KROLL, 2010, p. 90)

After seeing the preambles, participants saw an adjective. They were expected to repeat the preamble and complete the sentence with the adjective. The results indicate more elevated agreement mistakes for mismatched distributive-referent items than for mismatched single-referent items. While bilinguals were sensitive to grammatical and conceptual number in their L1, only higher proficiency participants were sensitive to conceptual number in the L2.

Even though a child's first language becomes automatic, adult L2 learners struggle to acquire some grammatical features, as "some aspects of L2 tend to fossilize" (JIANG, 2007, p. 5). According to Han (2010, 2013), morpheme acquisition depends upon conceptual restructuring, which is problematic for most learners[29]. He defends that while L2 learners have explicit knowledge about how plural (-s) is used, they do not have the necessary ability to use it automatically. This is the reason why we decided to study inflectional morphemes.

As a result, inflectional morpheme processing and production are very problematic for L2 learners and can have different outcomes if we consider the type of NP being analyzed and the proficiency levels of L2 learners. In our task, we investigate if the participants' proficiency level influences their results in detecting and correcting long-distance and local-distance agreement ungrammatical sentences.

2.1.6 Relational Morphology (RM)

Jackendoff & Audring (2016) propose a theory of linguistic representation and processing called Relational Morphology (RM). It aims to integrate phonology, morphology, syntax, and semantics. They believe that a language theory should regard the mental faculties while respecting their individualities. This theory focuses on morphology because it deals with linguistic organizational features that are not very significantly manifested in syntax.

RM is centered around the studies of Jackendoff (2002) and Culicover & Jackendoff (2005) about the Parallel Architecture, Booij's (2010) Construction Morphology, and Construction Grammar by Goldberg (1995). According to these theories, grammatical rules are treated as lexical items, which means that the grammar is integrated into the lexicon.

Most linguistic theories analyze a word on three levels: semantics, syntax, and phonology. The Parallel Architect conceives these levels independently but also linked to one another. This is what the authors call interface links; they are notated with indices that show the linkage from one level to the next. These indices mark the association lines between each category. We can see these

---

29 We cannot forget that speakers of some English varieties do not produce 3rd person singular morpheme (-s), such as in African-American Vernacular English (AAVE), as has been documented by Labov et al. since 1968. Yet, people alternate between standard English and informal English depending on the register.

mappings in (17) below:

(17) Mapping of the morphologically complex word *sheepish* (JACKENDOFF & AUDRING, 2016, p. 469):

Semantics: [SHEEP$_1$-LIKE; TIMID]$_2$
Morphosyntax: $[_A N_1 - aff_3]_2$
Phonology: /ʃip$_1$ ɪʃ$_3$/$_2$

The morphosyntax states that this word is constituted of a noun plus an affix. The coindexes are represented by the numbers above, which are relational links. Coindex 1 connects the syntactic category Noun to the meaning *SHEEP* and the phonology /ʃip/. Coindex 2 unites the semantics, morphosyntax, and phonology of the entire word. Coindex 3 connects the affix and its phonology /ɪʃ/. Besides, coindex 1 unites the significant chunks of *sheepish* and *sheep.*

Construction Grammar names a linguistic pattern such as (17) above a construction, whereas Construction Morphology names it a schema. Jackendoff & Audring follow the latter because they believe that it is declarative knowledge rather than procedural knowledge. There is a distinction between schemas and rules. The former follows the same format as words; the difference is between their structures, as schemas have variables and variable coindices. Once this is established, words and schemas can be stored in the same place in the mind and the theory. The lexicon and the grammar do not need to be seen as different constructs. Moreover, schemas seize the similarities among lexical items; therefore, they are easier to learn, store, and process. This is what they call the relational function of schemas.

Jackendoff & Audring claim that schemas are part of the declarative rather than the procedural representation since (18), as it is observable below, does turn input into output. The schema in (18) has structural layers, such as in (17). However, it contains open slots or variables instead of a lexical base:

(18) Schema to represent words that follow the same configuration as in "sheepish" in (2) (JACKENDOFF & AUDRING, 2016, p. 471)

Semantics: [X$_y$-LIKE]$_z$
Morphosyntax: $[_A N_y - aff_3]_z$
Phonology: /...$_y$ ɪʃ$_3$ /$_z$

In what concerns coindexation in (17), coindex 3 connects the morphosyntax affix to its phonology /ɪʃ/. Nonetheless, the variable coindex *z* connects to all the layers of the schema, and the variable coindex *y* connects to the internal variables.

Furthermore, schemas also have a generative function; they enable combinations. Take, for instance, the appearance of a word, such as *wug*. According to the authors, we would be able to construct a new adjective by attaching *–ish* to it. Similarly, schemas are also used for idioms – *miss the boat*.

Baayen; Dijkstra & Schreuder (1997) found data supporting the evidence that some Dutch nouns are stored rather than being created generatively. It is possible to prevent subcategorization conflict and save time by doing so. In a rule-based theory, this is such a complex issue since it is conceived that generated items cannot be part of the lexicon. This is what Langacker (1987) calls the "rule-list fallacy":

> 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 rules and lists (p. 42).

Langacker argues that English speakers learn, in addition to the general rule of plural formation, regular plural forms as fixed units[30]. Speakers acquire a large number of constructions, as fixed units, that are fully "analyzable and regular in formation". Moreover, a schema for the regular past tense has also stored forms. In this theory, it is specified through relational links instead of being generated. Schemas' function is not restricted to morphology. Therefore, one can also apply it to syntax. In the example (19) below, we can see the schema of English Verb Phrase (VP):

(19) [$_{VP}$ V – NP]

(JACKENDOFF & AUDRING, 2016, p. 472)

Productive schemas have both generative and relational functions, while nonproductive schemas have only the relational functional. However, schemas cannot have only the generative function. If a schema can be generated, it can also be stored. Additionally, the relational function is the base of the generative function. Consequently, the interactions inside the lexicon cannot be seen from a traditional generative perspective but rather from a relational perspective. The authors

30 This third option, in which both rules and lists can be used, is an alternative option for those trying to understand why L2 learners may not decompose regular verbs.

argue that Halle and Marantz's (1993) work about Distributed Morphology was an attempt to deal with nonproductive schemas considering generative rules.

A schema "motivates its instances" because word families motivate the other elements through an inheritance hierarchy. Items in a lower position inherit the structure from the ones that are in a higher position in the hierarchy as *sheepish* comes from *sheep*, such as in (20):

(20) Nsheep [A N-ish] Nchild

[A Nsheep ish] [A Nchild ish]

(JACKENDOFF & AUDRING, 2016, p. 474)

RM follows the notion of the impoverished entry theory that dictates that information in lexical items can only be derived from themselves. Hence, information from a lower node is also contained in a higher node. Sheepish in (21) has features from the other higher nodes:

(21) Nsheep [A N-ish]

[. . ; TIMID]

(JACKENDOFF & AUDRING, 2016, p. 475)

However, the authors do not believe that redundant details are deleted, such as suggested by the impoverished entry theory. A new complex word becomes part of one's inventory after the details have been fully mastered; otherwise, it is not possible to choose which schemas best fit and to which base the items will be attached. According to Booji (2017), the impoverished entry theory cannot cover the vastness of one's lexical memory. In his full-entry theory view, memorized complex words are fully specified and account for the acquisition of morphological schemas. Although one may acquire the English deverbal noun morpheme -er – such as in teacher and player – after a meaningful exposition, it is unlikely that one would erase the formulaic information once morphology is acquired. According to Hudson (2010), the mind finds similarities among items. Generalizations, such as these, are even used for creating new words.

Since economy does not seem suitable for a theory of lexical storage, Jackendoff & Audring adopt a full-entry theory approach, meaning that lexical items are entirely stored, even if there is redundancy. Moreover, Libben (2006) argues that redundancy enables a more robust computation.

During the process of language acquisition, children figure out the productive rules, that is, the productive schemas, from the input. Words and fixed phrases are stored in memory; nonetheless, storing items is not enough, and similarities among items have to be established in LTM. Similar items share sister links, also known as the constants in the schema, such as in sheep and sheepish. Therefore, analogies can be done by the prestructured basis in schemas.

Language knowledge has a network of linked nodes that involve semantics, morphosyntax, and phonology that cannot be seen apart. This means that the representational theory selects the components of the processing theory. The lexicon sustains spreading activation, but the interface and relational links regulate the spreading activation. The linkage extent determines the strength of spreading activation.

For Jackendoff & Audring, the processing is based on a network model. WM and LTM ought to be viewed as distinct components:

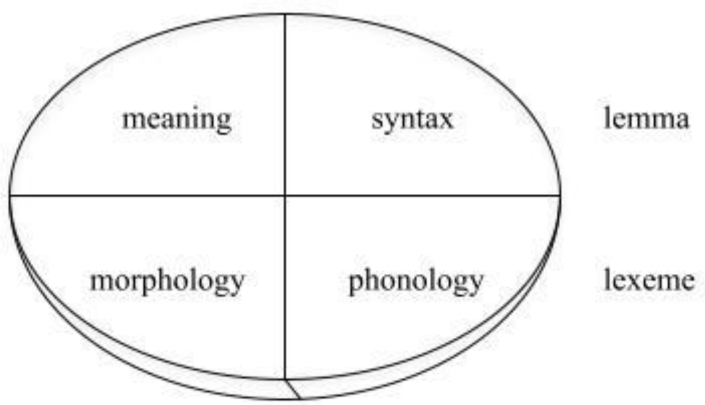
> LTM contains the lexical network of "knowledge of language" [...] It is in LTM that schemas fulfill their relational function, through their links to more fully specified items. In contrast, working memory (WM) is the functional component in which pieces of lexical items are assembled into larger structures, either to create an utterance (in production), or to analyze and parse an input (in comprehension) (JACKENDOFF & AUDRING, 2016, p. 480, 481).

Accordingly, productive schemas in RM have generative and relational functions. Therefore, if these schemas can be generated, they can also be stored as fully specified items. Having items stored as declarative knowledge enables us to explain why L2 learners cannot generate morphemes and decompose verbs but can store and use these items as fixed units (LANGACKER, 1987; JACKENDOFF & AUDRING, 2016; BOOJI, 2017). Computation becomes more robust because of the redundancy in generating and storing lexical items (LIBBEN, 2006). Learners try to find similarities among lexical items through links. Once similarities are established in LTM, these items reach a higher level of resting activation and are easier to be retrieved and produced. Thus, WM can only process and produce items with the support of LTM (JACKENDOFF & AUDRING, 2016). Our goal is to investigate if participants can recall, detect, and correct ungrammatical sentences without the 3rd person singular morpheme. We decided to adopt the RM because it accounts for morphological representation and processing considering aspects of LTM and WM.

2.1.7 The Role of Lexical Development in L2 Morpheme Acquisition

According to Levelt (1989, p. 182), "a speaker's mental lexicon is a repository of declarative knowledge about the words of his language". Each lexical entry has four features: i) meaning; ii) a set of syntactic properties; iii) morphological specification; and iv) phonological specification – syllable and accent structure. This means that an entry such as *eat* will have a meaning of "to put food in your mouth, chew it and swallow it[31]"; a syntactic category of verb, and as its arguments "subject" and "direct object"; a root form as *eat*, *eats* as its third-person inflection, *ate* as past-tense inflection and *eaten* as past-participle form; and *eat* has a monosyllabic structure. Following this view, an entry can be split up into two parts: lemma and lexeme. While the former contains semantic and syntactic information, the latter contains information about the morphology and the phonology of an entry[32], as depicted in FIG. 2.8:

Figure 2.8: The internal structure of a lexical entry adapted from Levelt (1989, p. 182):



Besides, the lexical hypothesis by Levelt (1989) assumes that the lexicon will always mediate the speaker's message in such a way that a particular syntactic form cannot be triggered on its own but rather mediated by a lexical item. Thus, the message will activate lexical items that will, in turn, make the grammatical encoder produce a certain syntactic structure considering the items' grammatical properties and order of activation. "The syntactic, morphological, and phonological properties of an activated lexical item trigger, in turn, the grammatical,

[31] Available at <https://www.oxfordlearnersdictionaries.com/definition/english/eat?q=eat>.

[32] Other theories also argue that words have grammatical information. Word Grammar sees words as concepts that include properties such as meaning, pronunciation, word-class, and language. Accordingly, in a speaker's mental repertoire, a word has "a meaning, a realization, a word-class, a syntactic valency, a language and a frequency" (HUDSON, 2010, p. 116).

morphological, and phonological encoding procedures underlying the generation of an utterance" (LEVELT, Op. cit., p. 181).

However, Jiang (2000) suggests some adaptations to the internal structure of a lexical entry in L2 acquisition. While learning a word in L1 implies learning new semantic and grammatical specifications at the same time, learning a word in L2 involves a different process. Therefore, it is more plausible to assume that the existing mappings will be triggered than assume that new concepts or new specifications will be created. This happens because "L2 words are learned mainly as formal entities because, here, the meaning is provided, either through association with L1 translation or by means of definition, rather than extracted or learned from the context by learners themselves" (JIANG, 2000, p. 50). Besides, when learning an L2, attention is drawn to the formal features, such as spelling and pronunciation. Therefore, semantic, syntactic, and morphological information is usually scarce.

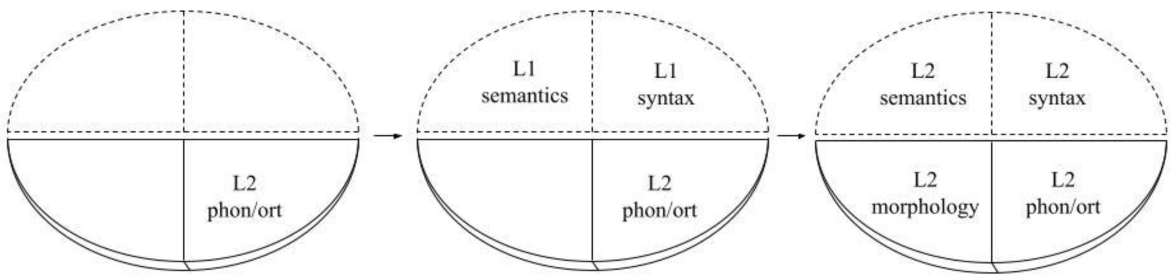
Nonetheless, it does not mean semantic, syntactic, and morphological information is not available. L2-L1 links activate semantic information and some syntactic information of L2 words, and learners have explicit knowledge about them:

> Under both circumstances, however, such semantic and grammatical information is not an integrative part of the mental lexicon. Instead, it is stored outside the mental lexicon, for example, as part of one's general memory or episodic memory and it can't be retrieved automatically in natural communication. In this sense, it is part of one's lexical knowledge, not one's lexical competence (JIANG, 2000, p. 51).

As a consequence, Jiang (Op. cit.) proposes a model, as represented in FIG 2.9, of lexical development in the L2. In the first stage, only formal specifications are present. The use of L2 words includes the activation of L2-L1 links. When the L2 learner's experience increases, these links between the L2 and the L1 translation become stronger. This entails the simultaneous activation of L2 word forms and L1 lemma information, which may eventually lead to a robust and straight connection between an L2 word and its L1 lemma translation. The second stage is reached when the entry includes L1 lemma information, which mediates L2 word processing. At this stage, L1 lemma information takes the space of the L2 lemma. That is, L2 words' semantic and syntactic information corresponding to L1 translation may be incorporated into the entry. In the final stage, L2 words' semantic, syntactic, and morphological specifications are integrated once this information is obtained through exposure.

Furthermore, these stages aim to explain the learning process of a word instead of the whole lexical knowledge. L2 learners can have words at different stages of development, but the majority of the words will define the stage in which they are. Nevertheless, there may also be some "gray areas"; as some words may be transitioning to another stage. This means that the learner can comprehend an L2 word without relying on its L1 translation but cannot produce it. Production involves vocabulary depth (SCHIMITT, 2014). Therefore, these stages do not have a "clear-cut":

Figure 2.9: "Lexical Development in L2: from the formal stage to the integration stage" by Jiang (2000, p. 54):



Other researchers have also investigated the difficulties faced by L2 learners and how resonance between L1 and L2 assists in L2 lexical development. MacWhinney (2017) investigated how entrenchment may influence L2 learning, especially for late learners. The Unified Competition Model (UCM) suggests that challenges faced by late learners such as entrenchment, transfer, overanalysis, and isolation could be counteracted by resonance, decoupling, chunking, and participation respectively. According to the author, entrenchment may help to explain L2 learners' difficulty with grammatical morphology learning and usage. Adults' learning process of lexical stems involves searching for very analytical forms, which ends up decreasing the chances of acquiring grammatical markers. Lexical learning in the L2 can be assisted by resonance as it supports new linguistic patterns through existing connections.

Jiang (2000) proposes an alternative to morphological information selection when used as explicit knowledge, namely, L2 learners' lexical knowledge. In this view, there are two steps followed: i) root selection and ii) morphologically suitable form selection. Therefore, an entry such as *leave* will have all of its morphological variations – *leave*, *leaves*, *left,* and *leaving* – stored individually in the L2 lexicon. He also suggests that it is reasonable to assume that these entries will have established links, which are made during the learning process. Thus, when one produces *they left*, the entry *leave* will be checked. Afterwards, morphological knowledge will be used in the selection of *left*. This morphological process is less automatic, which may result in morphological errors. When L2 users are engaged in natural communication, most of the time, these features are not the most crucial part of the message and are frequently disregarded by the mind.

As a result, we assume that there are links among linguistic contents – morphology, syntax, semantics, and phonology – (JACKENDOFF & AUDRING, 2016; HUDSON, 2010) and also between L2 and L1 (JIANG, 2000). Knowledge stored in LTM enables WM to assemble what will be processed and produced. LTM feeds information sent to WM because of a shared network of representation and processing (JACKENDOFF & AUDRING, 2016). In our task, participants have to recall, detect, and correct ungrammatical sentences without the 3rd person singular morpheme. In order to do so, they need their WMC to recall the sentences and their L2 lexical knowledge to correct these verbs, but this is only possible for those that have mastered the stages of lexical development. Explicit knowledge in the form of stored single verb forms can cause problems in lexical selection; thus, less proficient learners may have difficulties retrieving these forms. Many lower proficiency participants may rely on the connections between the L1 and L2 to retrieve and encode the stimuli. Some lexical items may have not been completely developed and, consequently, cannot be totally produced.

2.1.8 Theories Summary

The performance of L2 learners concerning the processing of inflectional morphemes and their production has already been investigated previously by many (see, for instance, JIANG, 2004, 2007; CARNEIRO, 2008, 2011; HOSHINO; DUSSIAS; KROLL, 2010; MOTA & BALTAZAR, 2015; JENSEN et al., 2019; OLIVEIRA; FONTOURA & SOUZA, 2020). The goal of this study is to investigate how BPE learners and natives behave in a task in which they have to recall, detect, and correct ungrammatical sentences without the 3rd person singular morpheme. We intend to enlighten whether 3rd person singular agreement can be detected and then corrected among non-natives and native groups (ERLAM, 2006). Besides, we check if L2 proficiency interferes with the non-native participants' performance in the task, as it is only at the end of lexical development that L2 learners have access to morphological information (JIANG, 2000). The shared network of processing and representation can also help explain how morphological information is processed and stored. Proficiency has had effects on L2 users' performance with inflectional morphology with long-distance and local-distance agreement sentences (HOSHINO; DUSSIAS & KROLL, 2010; JENSEN et al., 2019). The conditions of interest in our study introduced manipulations of local versus long-distance agreements. Our adapted sentence recall task demands both morphological and WM resources because participants have to identify and correct sentence violations and memorize and judge sequences of images in the intermediary memory task. Therefore, we analyze the role of WMC and proficiency in the participants' performance. WMC has been correlated with linguistic performance in some studies (FINARDI, 2009; AZEVEDO, 2012), but not in others (MOTA & BALTAZAR, 2015; FONTOURA 2018). WMC is also very important in SLA (WEN; MOTA; MCNEILL, 2013). Since participants have to store and manipulate the content of sentences presented to them, it is crucial to investigate if their WMC can impact their information storage and manipulation abilities.

2.2 Methods Background

Here we present the studies on which we base the Sentence Recall Task (SRT) and the Vocabulary Levels Test (VLT), which is the proficiency test that was used to rank participants.

2.2.1 Sentence Recall Task (SRT)

Sentence recall tasks demand that participants reconstitute sentences previously heard or read by either repeating or writing them; some of these request that participants reproduce sentences that exceed their working memory span. For the purpose of this task, participants have to recall and write sentences. Nevertheless, to do so, "the subjects must organize them in some manner; that is, since the sentences cannot simply parrot from short-term memory, the subjects must encode them through the use of some semantic, syntactic or other mnemonic device" (HAMAYAN; SAEGERT & LARUDEE, 1977, p. 86). In what concerns sentence recall tasks, we base our study on Souza (2014), Erlam (2006), Potter & Lombardi (1990, 1998), Lombardi & Potter (1992), Hamayan; Saegert & Larudee (1977), and Sachs (1967). These authors have already conducted studies with this paradigm to test English L2 learners' and natives' language knowledge and check whether they parrot the sentences presented to them or reconstruct the content of the sentences. Furthermore, Erlam (Op. cit.) also investigated inflectional morpheme sentences, which are critical to this study.

Erlam (Op. cit.) distinguishes between two types of sentence recall tasks, one that taps into implicit knowledge, and another that taps into explicit knowledge. The first type is reconstructive, while the second type relies on simple rote repetition. In this study, we focus on the first type because we intend to avoid simple rote repetition. According to the author, the main features in the first type are i) main focus on meaning instead of on form; ii) delay between stimulus presentation and repetition; iii) prompt correction of ungrammatical sentences; iv) no substantial connection between stimuli length and accomplished repetition; v) be completed under time pressure.

She conducted a study in which seventeen structures were investigated in an elicited imitation task. Syntactical and morphological structures were examined. Participants listened to both grammatical and ungrammatical sentences and were instructed to verbalize the sentences in the correct form. During the training, they received feedback about what they should have said. This method was selected to ensure that higher proficient participants would not simply parrot what was stored in WM. However, participants were not instructed that they would hear ungrammatical sentences. The sentences had between 8 and 18 syllables in length. After hearing the sentences, participants had to answer if they believed that the statement was either “true”, “not true” or “not sure” by circling it on a sheet. This design was adopted “to maximize the possibility that participants would focus on the meaning rather than the form of the sentences they heard” (ERLAM, 2006, p. 474). The topic of sentences varied, and they were constructed in a way that participants could give an opinion about them, such as in the examples (22) and (23) below:

(22) Princess Diana lov**ed** Prince Charles but divorc**ed** him – *Grammatical sentence with past tense.*

(23) *When man invent**ed** the motor car, life **change** for everyone – *Ungrammatical sentence with past tense.*

(ERLAM, 2006, p. 474)

Her participants consisted of a group of twenty English native speakers and ninety-five L2 learners. Native speakers repeated 97% of grammatical sentences accurately and corrected 91% of ungrammatical sentences. The results indicate that this elicited imitation test is a probable measure of implicit language knowledge. Furthermore, L2 learners repeated grammatical sentences correctly 61% and corrected ungrammatical sentences 35% of the time. Hence, as L2 learners corrected ungrammatical sentences, although to a lesser extent compared to grammatical sentences, it suggests that this kind of test does not tap into rote repetition. The key objectives of this study are testing if participants can detect and correct ungrammatical sentences; otherwise, participants would be simply recalling the ungrammatical sentences and relying on rote repetition.

Besides, three factors can be accounted for in reconstructive sentence recall tasks.The first factor relates to WMC regulated by stored knowledge. Both Baddeley (2003) and Cowan (1999) defend that working memory has an interface with long-term memory. Therefore, information could only be manipulated if it can be retrieved from the inventory. Baddeley; Gathercole & Papagno (1998) claim that the phonological loop is not only responsible for the acquisition of new words but also mediates syntactic knowledge; which may support the idea of internalized knowledge being manipulated.

Secondly, when reproducing a sentence, its meaning is preserved longer than its form. That is because the meaning can be presented in different ways, and WMC is limited to storing a surface representation. Levelt (1989) argues that the lexicon regulates grammatical encoding, that is, each word demands a type of syntactic structure. Thus, the form can be altered depending on the selected word. Potter & Lombardi (1990) discuss the difference between tasks that tap into verbatim memory and others that tap into long-term memory. While the former provides a surface representation of the sentence, the latter recalls only the meaning of the sentence, namely, its conceptual representation. In their experiments with synonym lures, the target words were substituted by the lures in 34% of the trials, whereas the other words were elicited highly accurately. Hence, immediate recall selects words recently activated. “Immediate recall of a sentence is like long-term recall in that it begins with a representation of the meaning of the material to be recalled” (POTTER & LOMBARDI, 1990, p. 646). In other words, it uses a regenerated conceptual representation[33] of the sentences with recently activated words. In addition, oral stimuli seem to intrude more than written stimuli. As a consequence, while the conceptual content stayed in the participants’ memory, the rest faded.

On the other hand, both studies of Lombardi & Potter (1992) and Potter & Lombardi (1998) argue that a syntactic representation can be primed by sentences previously heard or read, provided that the construction being reproduced is licensed. This means that not only lexical items can be kept in the surface representation but also a syntactic structure. A double-object dative sentence,

[33] This is also called the regeneration hypothesis. The conceptual and semantic representations are crucial to regenerate a structure, but both lexical items and syntactic constructions recently activated can play a role in what is being produced.

such as (24 a), can be primed by other preposition dative sentences resulting in (24 b). Therefore, learners will produce a structure that has been recently presented:

(24 a) The prompt secretary wrote her boss a message every week.

(24 b) The prompt secretary wrote a message to her boss every week.

(POTTER & LOMBARDI, 1998, p. 268).

When it comes to bilinguals, a structure from one language can be activated while they are processing another language. Souza (2014) investigated Portuguese-English bilinguals in sentence recall tasks using the induced movement construction, which is licensed in English, as in the sentences (25 a) and (26 a), but not in Portuguese, as in (25 b) and (26 b):

(25 a) The captain marched the soldiers to the camp.

(25 b) *O capitão marchou os soldados para o acampamento.

(26 a) The psychologists ran the rat through the maze.

(26 b) *Os psicólogos correram o rato pelo labirinto.

(SOUZA, 2014, p. 96).

Bilinguals performed the task in both languages while monolinguals just in their native language. Participants were instructed to read the sentences silently, orally recall the sentence, without any written aid, and judge whether a second sentence was the same they had just read and orally repeated or not. The results suggest that only the higher proficiency participants had access to the licensed construction from the L2 in the L1. This means that the accessibility to a specific type of construction from the L2 is modulated by proficiency. Therefore, this can be an indication that the two languages' linguistic representations are integrated when learners reach a certain proficiency level.

Moreover, Sachs (1967) conducted a study in order to test students' ability to retain syntactic and semantic memory. Participants had to rank if the sentences they heard were either identical or changed regarding their meaning and form. The results yield that "recognition for the form of a sentence declines much more rapidly than recognition memory for the meaning" (SACHS, 1967, p. 442). The form is just used for comprehension; afterwards, only the meaning is preserved. It is important to consider not only the meanings of isolated words but also the syntactic form, as the words can be combined in different forms and convey different meanings. The author defends that the meaning of a sentence is related to the deep structure (CHOMSKY, 1965) in the sense that it is used for recalling. However, the surface structure is lost whenever it is not important for the sentence's meaning. She attributed the fact that some students could remember the form after 80-160 syllables to chance. On the other hand, subtle changes in meaning were noticed by the majority. Immediate recall of sentences was very high for all test types.

The last factor that Erlam (2006) holds responsible for reconstructive recalling is the ability to correct ungrammatical sentences spontaneously. Hamayan; Saegert & Larudee (1977) compared the performance of three groups in elicited imitation tests that aimed to exhaust the memory span. After the target sentence (9 syllables in length), an explanatory sentence (5 syllables in length) was placed to make sure participants were not using their WM span, as in example (27):

(27) The boy who was running fell down. He broke his arm.

(HAMAYAN; SAEGERT & LARUDEE, Op. cit., p. 88)

In addition, a line of drawings was presented after the sentences so that the understanding was ensured. Participants received the instruction to repeat the sentences exactly as they had heard. Both grammatical and ungrammatical sentences were presented to participants. All of the groups were Arabic-English learners. The first group comprised 8-year-olds, the second one 11-year-olds, and the third one adult learners. The results suggest that the younger group had more problems normalizing type A (violations on conjunction, complement, and number) than type B sentences (violations on negative wh-question, relative clause, auxiliary verb, and adjective). There was only a moderate difference for the intermediate group and no significant difference for the adults. Therefore, type A sentence normalization increased as a function of age. An interaction of sentence type by groups was found.

Accordingly, this study intends to implement an adapted version of a sentence recall task to be carried out in a typed version instead of in a spoken version because of the imposed limitation of carrying out an online experiment during the pandemic. We expect that this type of task will enable us to examine how L2 learners behave in a task that demands reproducing grammatical sentences with the 3rd person singular morpheme;, detecting and correcting ungrammatical sentences without the 3rd person singular morpheme, and memorizing and judging the sequence of letters and images. Our task has an intermediary memory task between the sentence exhibition and the sentence recall to test participants' WMC (adapted from Kane et al. (2007) and

Fontoura(2018)). Following Erlam (2006), we investigate if participants can reconstruct the content of the sentences and correct the violations they come across. We expect that the proficiency level would influence the ability to correct these violations. The author claims that the main features in reconstructive sentence recall tasks are (i) focusing on meaning rather than on form; (ii) having a delay between sentence presentation; (iii) promptly correcting ungrammatical sentences; (iv) having no significant correlation between stimuli length and successful recall; and (v) completing the task under time pressure. We tried to control as many factors as possible to have a reconstructive sentence recall task; however, there were some limitations to these factors in our task. Participants saw the sentence for 5,000 msec, conducted the images judgment, and then typed the sentence in up to 100,000 msec. This format was intended to avoid a rote repetition and make participants focus on the meaning. Our task had written stimuli and participants typed the sentences recalled; therefore, we could not check if participants promptly corrected ungrammatical sentences. Since our task was recalled in written format, participants could access their explicit knowledge and monitor what they wrote. Nevertheless, only higher proficiency participants could retrieve and use this type of knowledge. Besides, we controlled the stimuli length because it was exhibited in a limited time frame.

2.2.2 Vocabulary Levels Test (VLT)

Proficiency level can influence the results of participants in recalling, detecting, and correcting ungrammatical sentences (ERLAM, 2006). Therefore, we divide our bilingual participants into lower proficiency and higher proficiency to check if the two groups' performance could be related to their proficiency level.

Alderson (2005) claims that it is possible to correlate language performance with vocabulary size. This motivated us to adopt the Vocabulary Levels Test (VLT) to measure the proficiency of our participants. Souza & Silva (2015) managed to validate VLT in comparison to the Oxford Placement Test (OPT) for the Brazilian university students' population of English L2 learners. Furthermore, the vocabulary size tested on VLT corresponds to lemma rather than isolated words; therefore, it does not consider words isolated. The level of proficiency in which participants were categorized matches the Common European Framework of Reference for Languages (CEFR):

> 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).

The original VLT format consisted of a matching task in which participants had to choose three among six options[34], as represented in FIG. 2.1:

Figure 3.1: 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	

Participants needed to score at least 12 items out of 18 so that they could advance to the next section (NATION, 1990). In order to be considered high proficient, participants needed to reach level 5 as Souza; Duarte & Berg (2015) could not find any significant effect for level 4. The authors attribute this result due to the fact that level 4 consists of Latin-originated words related to academic and scientific vocabulary, which many of them are cognates to Portuguese.

Although VLT had no time restriction in Nation (1990), Laufer & Nation (2001) found results that yield a correlation between RT and vocabulary size. As a result, Souza; Duarte & Berg (2015) imposed a time constraint of 10 minutes to complete VLT. Moreover, Silva (2016) believes that this time restriction suits the test for the population under investigation, as level 4 was disregarded.

2.2.3 Methods Background Summary

The SRT will enable us to check the three proficiency groups' ability to (a) process, store and recall grammatical sentences with 3rd person singular morpheme; (b) detect and correct long-

[34] In this version, participants performed the test on paper. However, it will be performed online such as in Fontoura (2018).

distance and local-distance ungrammatical sentences without the 3rd person singular morpheme; and (c) recall and judge sequences of images and letters. We will divide participants into lower and higher proficiency depending on their performance on VLT. Proficiency can impact the participants' performance when they (i) recall and correct the sentences; and (ii) memorize and judge the sequences of images. We also investigate if WMC can impact the learners' performance.

CHAPTER 3

MATERIALS AND METHODS

In this chapter, we describe the materials and methods that are used to implement this study, which include the experiment configuration and the stimuli used. These materials and methods will help us answer our five research questions:

a. Is there any distinction in the performance of higher proficiency, lower proficiency, and native participants when contrasting their performance in reproducing a grammatical control condition with sentences inserted in a context where the 3rd person singular morpheme (-s) is expected?

b. Is there any difference in the performance of the three groups when we compare their performance in detecting and correcting four target ungrammatical conditions with the long-distance agreement – animate and inanimate – and local-distance agreement – name and pronoun – without the 3rd person singular morpheme (-s)?

c. Does WMC influence the groups' performance in reproducing grammatical sentences with 3rd person singular morpheme (-s)?

d. Does proficiency level result in better detection and correction of ungrammatical sentences without 3rd person singular morpheme (-s)?

e. Do the three groups have a similar performance in image and considering proficiency and working memory capacity?

And test our five hypotheses:

a. Grammatical sentences with 3rd person singular morpheme are more accurately reproduced than correcting ungrammatical sentences without the 3rd person singular morpheme within the groups.

b. Long-distance agreement ungrammatical sentences without the 3rd person singular morpheme are harder to detect and correct than the local-distance agreement ungrammatical sentences without the 3rd person singular morpheme within the groups.

c. There is greater demand for working memory capacity for the lower proficiency group, affecting the reproduction of grammatical sentences with the 3rd person singular morpheme compared to the other groups.

d. The proficiency level influences the participants' ability to detect and correct ungrammatical sentences without the 3rd person singular morpheme between groups.

e. Proficiency does not impact image recall and judgment between the groups.

3.1 Structure of the Study

The data for this study was gathered in two experimental sessions with the support of two platforms. The experiment and the survey were administered online on Psytoolkit (for further information check Stoet (2010, 2017) and Appendix 6) and VLT on classmark (Appendix 5)[35]. The organization of this study was carried out as follows in table 3.1:

Table 3.1: Structure of the study:

Experiment	online
Vocabulary Levels Test (VLT)	

3.1.1 Experiment

Firstly, the welcome screen introduced the nature of the task and invited participants to carry it out (check Appendix 2 for the instructions of the task[36]). Participants answered some questions before starting the experiment. Some of these questions differed for native and L2 participants (check Appendix 3). After the experiment was conducted, they answered questions about their typing abilities (check Appendix 3). The objective of these questions was to ensure that participants fit all the requirements that we set and get to know more about their backgrounds.

In the experimental part, instructions were presented regarding what was expected to be done. Before each sequence began, a message appeared on the screen for 2,000 milliseconds (msec)

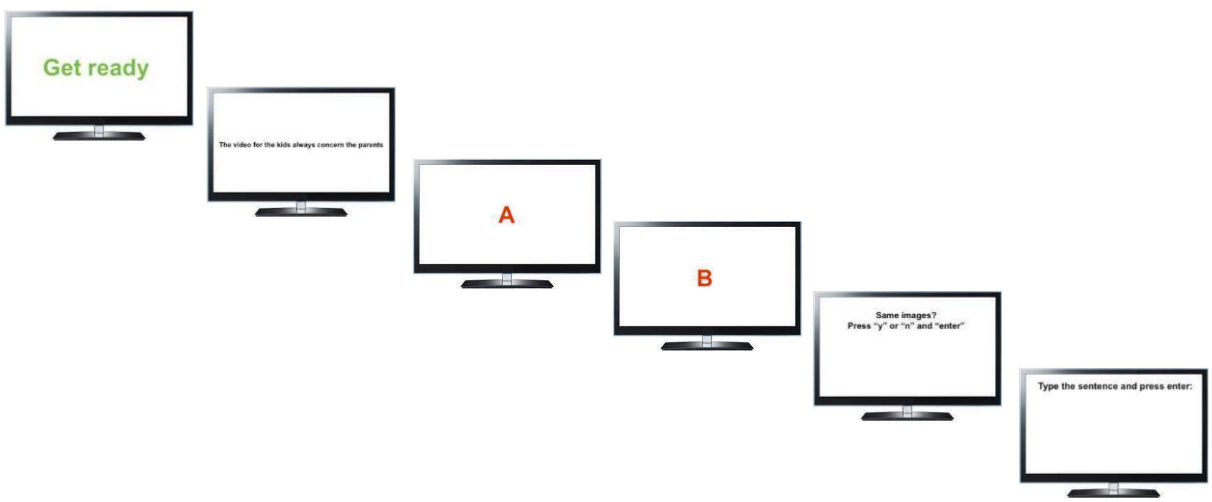
[35]Available at < www.classmarker.com>.

[36] Instructions to the task were provided in the participants' native language.

for participants to get ready. It is important to note that there was a delay of 1,000 msec between each display. Afterwards, the sentence was presented for 5,000 msec[37], followed by the first image and then the second presented for 2,000 msec each. The next screen asked participants to type whether the images were the same or not. Finally, participants had to enter the memorized sentence, as depicted in FIG. 3.2.

Participants received feedback concerning their performance on the image judgment in the training session. They saw whether the pressed key was correct, incorrect, or had to be faster. A timeout of 10,000 msec was established for this part. On the other hand, they did not receive any feedback on the sentence they entered. Yet, there was a timeout of 100,000 msec to enter the sentence.

Figure 3.2: Representation of the task[38]



Before starting the experiment, participants conducted a training session following the same parameters as the actual trial, as in table 3.2:

[37] In Souza & Silva (2015), participants had up to 6,000 msec to read and judge the sentences with up to 40 characters excluding spaces.
[38] All instructions, sentence display, and letters used Arial font.

Table 3.2: Training Session

Sentence	Image 1	Image 2
The illustrations in the manual can help	couch_1.jpg	couch_2.jpg[39]
They workt many hours	letter_X_1.jpg	letter_X_2.jpg[40]
We lived in USA	birds.jpg	birds.jpg[41]
The label on the bottles are always black	letter_q_lower.jpg	letter_q_lower.jpg[42]

3.2 Pilot Version

Before collecting data for this study, we decided to conduct a pilot version with five BPE learners. They were submitted to the experiment and then to VLT. In this version, participants received no explicit instruction to correct sentences with violations.

3.2.1 Pilot Sentences

For the pilot version, we had four ungrammatical sentence conditions and four grammatical sentence conditions. The ungrammatical conditions (table 3.3, 3.4, 3.5, and 3.6) were considered the main target, whereas the grammatical conditions (table 3.7, 3.8, 3.9, and 3.10) were the controls. There were six sentences for each ungrammatical condition and four for each grammatical condition. The first and the second conditions of the ungrammatical conditions were long-distance agreements and changed concerning the animacy of their NPs, table 3.3 contains animate nouns and table 3.4 inanimate nouns in the nucleus of the NP in the subject position[43]:

---

[39] Mismatching images.
[40] Mismatching letter colors.
[41] Matching images.
[42] Matching letters.
[43] Information about the words selected and the sentences' formation is provided below in section 3.5.

Table 3.3: Ungrammatical: The animate noun + with + the inanimate noun + frequency adverb + verb + complement

Animate/Inanimate nouns	frequency adverb	Verb	Complement	Syllables	Characters with space
The **engineer** with the **tools**	usually	***build***	many[44] houses	16	53
The **girl** with the **dolls**	usually	***wear***	sunglasses	13	47
The **judge** with the **papers**	usually	***show***	results	13	46
The **teenager** with the **keys**	usually	***open***	doors	14	45
The **farmer** with the **crops**	usually	***produce***	fruits	13	48
The **boy** with the **hotels**	usually	***draw***	attention	14	49

Table 3.4: Ungrammatical: The inanimate noun + for + the animate noun + frequency adverb + verb + complement

Inanimate/Animate nouns	frequency adverb	Verb	Complement	Syllables	Characters with space
The **uniform** for the **soldiers**	always	***require***	planning	14	52
The **video** for the **kids**	always	***concern***	the parents	13	49
The **gift** for the **artists**	always	***display***	many pictures	14	53
The **van** for the **managers**	always	***charge***	a lot of cash	14	52
The **machine** for the **teachers**	always	***print***	many tests	13	52
The **menu** for the **waiters**	always	***describe***	the options	14	51

[44] The word "many" was not supposed to be in this version, but it failed to be excluded from the script on Psytoolkit.

The third and fourth conditions are local-distance agreement with names (table 3.5) and pronouns (table 3.6):

Table 3.5: Ungrammatical: Name + frequency adverb + verb + complement

Name	frequency adverb	Verb	Complement	Syllables	Characters with space
Sarah	always	***eat***	breakfast	7	26
Brian	always	***read***	nice books	7	28
Lucy	always	***see***	the friends	7	27
James	always	***clean***	the room	6	28
John	always	***feed***	the bird	6	25
Kate	always	***send***	flowers	6	24

Table 3.6: Ungrammatical: Pronoun + frequency adverb + verb + complement

Pronoun	frequency adverb	Verb	Complement	Syllables	Characters with space
He	usually	***paint***	walls	7	22
She	usually	***buy***	food	7	20
He	usually	***call***	the boss	8	24
She	usually	***pay***	the rent	8	24
He	usually	***drink***	juice	7	22
She	usually	***tell***	jokes	7	22

The grammatical conditions followed the same pattern. The long-distance conditions had animate nouns, as it is illustrated in table 3.7, and inanimate nouns, as in 3.8, in the NP in the subject position:

Table 3.7: Grammatical: The animate noun + with + the inanimate noun + frequency adverb + verb + complement

Animate/Inanimate	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
The **student** with the **pens**	usually	**makes**	a mess	13	46
The **chef** with the **plates**	usually	**serves**	pasta	13	45
The **vet** with the **cages**	usually	**holds**	animals	14	44
The **actor** with the **hats**	usually	**receives**	money	14	46

Table 3.8: Grammatical: The inanimate noun + for + the animate noun + frequency adverb + verb + complement

Inanimate/Animate	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
The **coffee** for the **directors**	always	**needs**	sugar	13	47
The **ring** for the **models**	always	**costs**	a great deal	12	49
The **radio** for the **singers**	always	**plays**	the hits	12	47
The **tie** for the **clowns**	always	**presents**	many colors	13	50

The grammatical local-distance agreement conditions also had names, such as in table 3.9, and pronouns, as in table 3.10, in the subject position:

Table 3.9: Grammatical: Name + frequency adverb + verb + complement

Name	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
Austin	always	**uses**	drugs	7	24
Alice	always	**sells**	products	7	27
Noah	always	**asks**	questions	7	28
Anne	always	**bakes**	good pies	7	27

Table 3.10: Grammatical: Pronoun + frequency adverb + verb + complement

Pronoun	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
She	usually	**speaks**	French	7	25
He	usually	**writes**	letters	8	25
She	usually	**rides**	a bike	8	24
He	usually	**trains**	dogs	7	22

We also included a fifth ungrammatical condition. Our goal was to check whether participants would preserve the violations they came across or correct simpler violations, compared to 3rd person singular morpheme violations. There were six sentences in this condition. The first three sentences follow a different NP pattern in the subject position from the last three sentences (table 3.11):

Table 3.11: Cut-off condition: NP + (auxiliary +) verb + complement

NP	**(auxiliary +) Verb**	**Complement**	**Syllables**	**Characters with space**
She	***are having***	problems	6	23
He	***could studied***	English	6	24
She	***losed***	the wallet	5	20
The **doctor** with the **toys**	***earnde***	a good salary	12	45
The **hunter** with the **guns**	***will to kill***	the duck	11	46
The **pilot** with the **flags**	***have driven***	a red car	12	46

3.2.2 Pilot Results[45]

We gathered a small sample of five BPE learners: three of them were higher proficiency, VLT 5, and the other two were lower proficiency, VLT 4 and VLT 3. The data we present here is a combination of their outcomes. We analyzed the number of correct and incorrect sentences in the cut-off, grammatical, and ungrammatical conditions. The sentences in the grammatical condition had to be reproduced, but the sentences needed to be corrected in the cut-off and the ungrammatical conditions. In table 3.12 below, we can observe that participants corrected only 0.1 of cut-off sentences. Furthermore, only 0.133 of ungrammatical sentences were corrected. However, participants correctly reproduced 0.825 of the grammatical sentences.

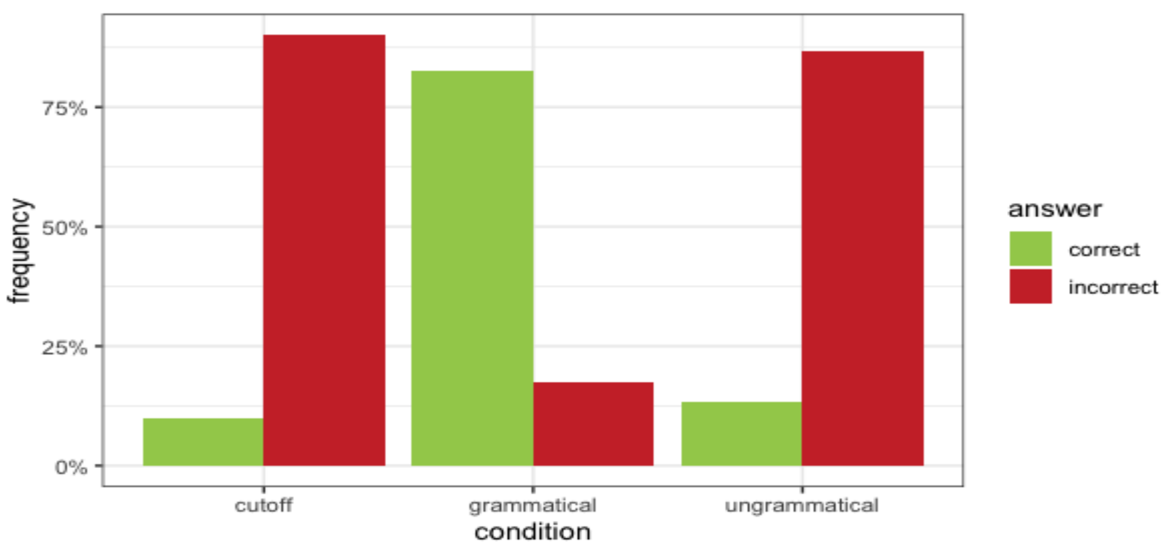
[45] All the data were analyzed using R, based on Godoy (2019).

Table 3.12: Correct and Incorrect Answers in the Sentence Conditions in the Pilot

condition	answer	n (items X participants)	frequency
cut-off	correct	3	0.1
cut-off	incorrect	27	0.9
grammatical	correct	66	0.825
grammatical	incorrect	14	0.175
ungrammatical	correct	16	0.133
ungrammatical	incorrect	103	0.867

Graph 3.1 represents the frequency of correct and incorrect answers in the sentence conditions:

Graph 3.1: Correct and Incorrect Answers in the Sentence Conditions in the Pilot



Our results indicate that participants preserved most of the violations on ungrammatical sentences and cut-off sentences and correctly reproduced the majority of grammatical sentences.

Considering this, we decided to reformulate the sentences[46] and provide explicit instruction for participants to correct the sentences. Erlam (2006) advocates that explicit instruction to correct sentences avoids highly proficient participants simply using rote repetition. Our main goal is to check if participants can detect and correct the violations they come across. We are also testing if their performance changes between different proficiency groups. Besides, we received feedback from participants that the task was too long and exhausting. Therefore, we decided to reduce the grammatical conditions and include a typo condition and a modal condition to have a better understanding of our previous condition named cut-off condition.

We grouped the participants' judgment of images and letters into correct and incorrect and all into the label image. In table 3.13, we can observe that participants did not get right 0.0696 of the sequences of images and letters:

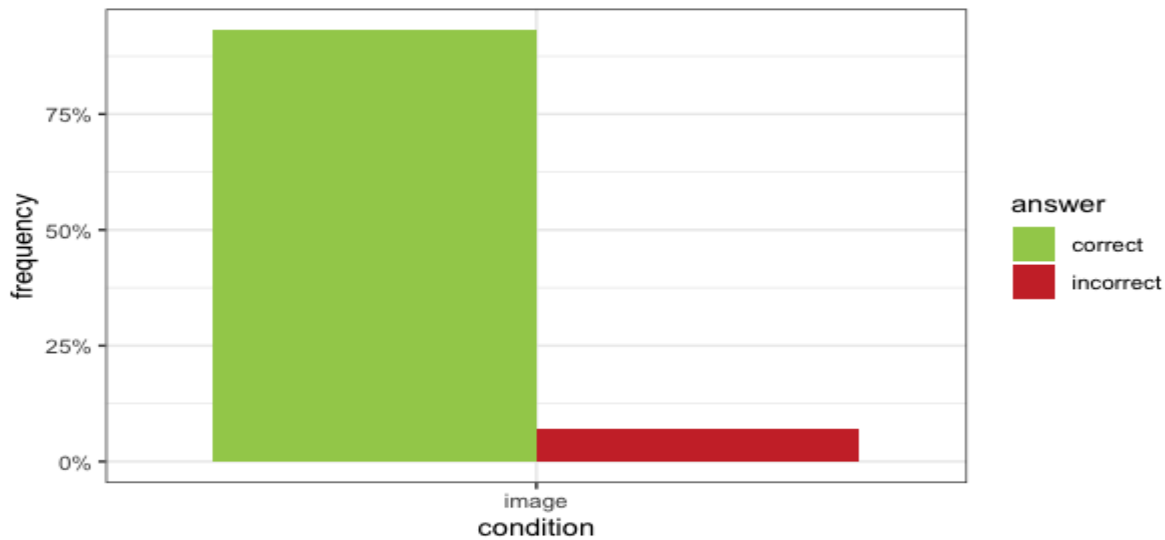
Table 3.13: Correct and Incorrect Answers in the Images Judgment in the Pilot

answer	**n (items X participants)**	**frequency**
correct	214	0.930
incorrect	16	0.0696

Graph 3.2 represents the frequency of correct and incorrect images sequences that participants entered in the judgment:

46 We had initially planned to have a condition with the 3rd person singular morpheme oversuppliance, such as done by Jensen et al. (2019). However, we had to adapt it because of the restraint imposed by collecting data online during pandemic times. Having more conditions would consume more time and concentration from participants.

Graph 3.2: Correct and Incorrect Answers in the Images Judgment in the Pilot



The images and letters presented few mistakes in the judgment and did not seem to be problematic; thus, we decided to keep the same format.

3.3 Restructured Target Sentences

Inflectional morphemes have long been problematic for L2 learners (SLABAKOVA, 2014; JIANG, 2004, 2007; CARNEIRO, 2008, 2011; FONTOURA, 2018), but 3rd person singular morpheme seems to be a special challenge for English L2 learners because of its lack of perceptual salience (GOLDSCHNEIDER & DEKEYSER, 2001; N. ELLIS, 2017). Moreover, conceptual agreement production is a challenge in the L1 and it can be even harder in the L2, being mediated by the proficiency level (HOSHINO; DUSSIAS & KROLL, 2010). Different types of NPs on the subject position can recruit different resources from participants, such as on local and long-distance agreements, which also interact with proficiency level (JENSEN et al., 2019).

Unlike English, Portuguese has no overt subject-verb (SV) singular agreement morpheme for third-person, as exemplified in the sentence (28 a), having an overt morpheme for SV plural agreement, as in (28 b). Therefore, the attentional cues that Brazilian learners of English may search are different from the ones natives direct their focus[47]:

(28 a) Ele canta bem.
He sing(Ø) well.
'He sings well'.

(28 b) Eles cantam bem.
They sing(3rd P PL) well.
'They sing well'.

Considering that the sentence was displayed for 5,000 msec, each sentence had its number of syllables[48] and characters[49] controlled. The number of syllables did not vary more than one and the characters no more than four for each condition. Although the sentences were presented and recalled in written form[50], participants could try to verbalize them to memorize; thus, having unbalanced syllables and characters could demand more attentional resources from participants, which we accounted for inside each condition. In addition, words were selected according to a rank on COCA[51]. They needed to be up to the 3,000th position because of the criterion specified on VLT, namely, level two's participants know the most frequent 3,000[52] words. Since we were considering participants starting on level two, we expected they would be familiar with the words in the sentences.

All of our target sentences followed a pattern of NP + frequency adverb + verb + complement[53]. In what concerns the NPs in the subject position, all animate nouns had to be human, but we also have a condition with inanimate nouns in order to test if there is any significant difference between these conditions. Nouns had to be regular in their plural form and could not end

---

47 Oliveira, Fontoura & Souza (2020) found evidence indicating that oversuppliance of 3rd person singular morpheme is more easily spotted than the morpheme omission.

48 Using the website Syllable Counter <https://syllablecounter.net/count>.

49 Including spaces.

50 We intended to conduct a sentence recall task in an oral format, but once more we had to adapt the task format to be performed remotely.

51 Available at <https://www.english-corpora.org/coca/>.

52 This criterion considers word roots; therefore, we checked the roots in inflected and derived words.

53 Our starting point for the long-distance NPs were the ones in the study of Hoshino; Dussias & Kroll (2010). However, we adapted them to standardize the prepositions in the conditions and consider the frequency of words in COCA. We also avoided NPs that could evoke a possessive ('s) placement by participants. After the adaptations, the NPs became different from the originals.

in a sibilant sound, as the plural form imposes another syllable to these forms (e.g. brush: /brʌʃ/, brushes: /brʌʃəz/). Verbs could not end in a sibilant for the same reason described for the nouns when they assume their 3rd person singular agreement form. Both nouns and verbs could not end in -y because whenever the first takes the plural form and the latter the 3rd person singular agreement form, they change their base written form (e.g. study and studies). Besides, the frequency adverb selected was always. There are no prepositional complements and no possessive adjectives on object position. We believe that the latter could help solve the conceptual agreement problem in long-distance sentences. Thus, we decided not to include them.

We selected local and long-distance agreements as the target items for this study because we aim to test the variation that may arise from lower proficiency, higher proficiency, and native groups in these contexts. Our goal is to examine if local-distance conditions are easier to correct than long-distance conditions for the lower proficiency group. We have four ungrammatical target conditions without the 3rd person singular morpheme – two long-distance agreements and two local-distance agreements – to check if different types of NPs can help participants to have a better perception of the violations and, consequently, yield more corrections. We set animate (table 3.14) and inanimate nouns (table 3.15) in the subject position in the long-distance agreement conditions:

Table 3.14: Ungrammatical Animate: The animate noun + with + the inanimate noun + frequency adverb + verb + complement

Animate/Inanimate nouns	frequency adverb	Verb	Complement	Syllables	Characters with space
The **engineer** with the **tools**	always	***build***	big houses	13	51
The **girl** with the **toys**	always	***wear***	beautiful clothes	12	52
The **judge** with the **papers**	always	***show***	good results	12	50
The **grandmother** with the **keys**	always	***open***	blue doors	13	52
The **sister** with the **hotels**	always	***draw***	attention	12	48
The **boy** with the **cameras**	always	***take***	pictures	13	51

Table 3.15: Ungrammatical Inanimate: The inanimate noun + for + the animate noun + frequency adverb + verb + complement

Inanimate/Animate nouns	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
The **video** for the **kids**	always	***concern***	the parents	13	49
The **gift** for the **artists**	always	***display***	many photos	14	51
The **boat** for the **workers**	always	***charge***	a lot of cash	13	51
The **machine** for the **teachers**	always	***print***	many tests	13	52
The **computer** for the **employees**	always	***store***	data	13	48
The **coffee** for the **students**	always	***need***	sugar	14	51

In the local-distance agreement conditions, we have names (table 3.16) and pronouns (table 3.17) in the subject position:

Table 3.16: Ungrammatical Name: Name + frequency adverb + verb + complement

Name	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
Sarah	always	***eat***	breakfast	7	26
Brian	always	***read***	nice books	7	28
Lucy	always	***see***	the friends	7	27
James	always	***clean***	the room	6	28
John	always	***learn***	new things	6	28
Kate	always	***hire***	young people	7	29

Table 3.17: Ungrammatical Pronoun: Pronoun + frequency adverb + verb + complement

Pronoun	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
He	always	***paint***	large walls	6	27
She	always	***buy***	healthy food	7	27
He	always	***call***	the police	7	25
She	always	***complete***	the task	7	28
He	always	***drink***	orange juice	7	28
She	always	***answer***	the phone	7	27

On the other hand, we have three control conditions. The main purpose is to check whether the three groups will preserve the grammaticality, correct the typos, and supply the infinitive mark. All the NPs in the subject position are names. The grammatical condition (table 3.18) and the typo condition (table 3.19) have a similar pattern presented in the local-distance ungrammatical condition with names:

Table 3.18: Grammatical: Name + frequency adverb + verb + complement

Name	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
George	always	**gets**	good grades	7	30
Alice	always	**sells**	products	7	27
Noah	always	**asks**	questions	6	26
Anne	always	**meets**	the brother	7	29
Austin	always	**speaks**	French	6	27
Jane	always	**writes**	letters	6	26

The typo condition presents its violations on the verbs. We decided to invert the last two letters of the verbs' base form and keep the 3rd person singular agreement mark (-s) at the end, as it is possible to observe in table 3.19:

Table 3.19: Typo: Name +frequency adverb + verb + complement

Name	**frequency adverb**	**Verb**	**Complement**	**Syllables**	**Characters with space**
Bob	always	***vistis***	the uncle	8	27
Meg	always	***drievs***	a red car	7	27
Max	always	***solevs***	the problem	8	31
Emma	always	***cokos***	nice meals	8	28
Henry	always	***statrs***	the fight	7	29
Kim	always	***hepls***	the father	7	27

Finally, the modal condition (table 3.20) presents an oversuppliance of the infinitive mark (to):

Table 3.20: Modal: Name + will to + verb + complement

Name	**Verb**	**Complement**	**Syllables**	**Characters with space**
Claire	***will to make***	a mess	6	26
Joe	***will to play***	football	6	26
Eve	***will to ride***	a bike	6	23
Fred	***will to pay***	the bill	6	25
Liam	***will to kill***	the bird	6	26
Sam	***will to tell***	jokes	5	22

3.4 Images

The intermediary task displays both images[54] and letters[55]. The aim is to have a non-linguistic task similar to what was done previously with letters by Kane et al. (2007) in a two-back and a three-back task, numbers by McDonald (2006, 2008), and letters and numbers by Mota & Baltazar (2015) and Fontoura (2018). We believe that having a linguistic task such as a sentence recall and another task that involves non-linguistic stimuli can tap into different memory components.

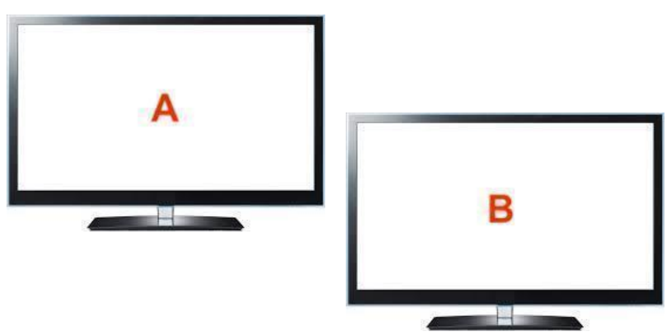
Therefore, participants see a sequence of two letters or two images and have to decide whether they match or mismatch. Fifty percent of the sequences match and the other fifty mismatch. The mismatching sub-conditions consist of uppercase letters that mismatch in letter sequence (letter_A/ letter_B) and also in color (letter_C_1/ letter_C_2) and also a sequence of images that mismatch (armchair_1/armchair_2), as exemplified in table 3.21:

Table 3.21: Mismatching Letters and Images

Uppercase letters mismatching letters	letter_A.jpg	letter_B.jpg
Uppercase letters mismatching colors	letter_C_1.jpg	letter_C_2.jpg
Images mismatching	armchair_1.jpg	armchair_2.jpg

Figure 3.3 illustrates the mismatching in letters sequence, figure 3.4 the mismatching in letters colors, and figure 3.5 the mismatching in images:

Figure 3.3: Uppercase letters mismatching letters



[54] All images were downloaded from gettyimages <https://www.gettyimages.com.br/>.

[55] All letters were created on a docx file and saved.

Figure 3.4: Uppercase letters mismatching colors

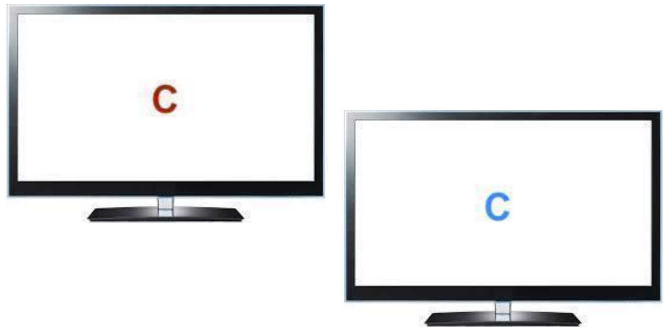


Figure 3.5: Images mismatching

The matching sub-conditions consist of lowercase letters that match both in letter sequence and color (letter_a_lower/letter_a_lower/) and images that are exactly the same (balloon/balloon), as we can see in table 3.22:

Table 3.22: Matching Letters and Images

Lowercase letters matching	letter_a_lower.jpg	letter_a_lower.jpg
Images matching	balloon.jpg	balloon.jpg

Figure 3.6 shows the matching in letters sequence and letters colors, and figure 3.7 as the images match:

Figure 3.6: Lowercase letters matching

Figure 3.7: Images matching

It is important to note that the images are complex, in both matching and mismatching conditions (for the complete list, check Appendix 1). Our goal is to allocate different types of attentional resources since attention is selective (COWAN, 1999).

Even though the uppercase letters and the lowercase letters had sequences using the same letters, no sequence of letters with different cases corresponds (no coincidence of the type: letter_C_1/letter_C_2 and letter_lower_c/ letter_lower_c).

3.5 Participants

In this study, there are three groups[56]: lower proficiency, higher proficiency, and natives. The lower proficiency group had to reach at least level 2 and the higher proficiency group level 5

[56] Participants had access to the free and informed consent term (Appendix 4).

on VLT (Appendix 5)[57]. All our bilingual BPE were living in Brazil during the test conduction and the native speakers were living in countries where English was the most dominant language. Having three proficiency groups enables us to contrast if the proficiency level influenced participants' performance on reproducing grammatical sentences, detecting and correcting ungrammatical sentences, and judging sequences of images.

We decided to limit the participants' age from 18 to 35-year-old. According to Salthouse (2009), even healthy educated adults can present cognitive decline in their 20s or 30s. Therefore, having a wider age sample could affect the results of our task, especially because it expects that participants store and manipulate images and language content. One can have their WMC impaired with age. Besides, education plays a major role in our study; thus, participants need to have, at least, completed high school. As we are testing the ability to distinguish one letter from another based on color, individuals that have any type of color blindness could not take part in this study.

[57] In this version, participants had eleven minutes to complete VLT, including the explanation part. We believe that this online format did not enable us to solve any doubt that could arise from participants; thus, we decided to give them an extra minute to get acquainted with the task.

CHAPTER 4

RESULTS

In this chapter, we discuss the findings of our study concerning the experimental conditions – the control grammatical condition, the target ungrammatical conditions: animate, inanimate, noun, and pronoun – and the three proficiency groups – lower, higher and native. The results are compared within the groups and then between the groups.

4.1 Participants

We collected data from 59 bilinguals and 15 natives. Since bilinguals had to undergo VLT and reach at least level 2, those that failed to complete the test or score level 2 were eliminated from the study. Furthermore, in order to be suitable for this study, participants could not leave more than 20% of sentences blank or incomplete (no more than 8 out of 42 sentences). Participants that missed or got wrong more than 20% of the sequences of images and letters were also excluded (more than 8 of 42 sequences).

It is important to note that six BPE participants were eliminated because they failed to reach at least level 2 on VLT, and ten did not find time to complete VLT during the months we were gathering data. One BPE participant had to be eliminated because he had color blindness. Two native speakers had to be excluded because they got wrong or missed more than eight sequences of images, and one was excluded because left more than eight sentences blank. After this screening, we were left with eight VLT 2, seven VLT 3, six VLT 4, twenty VLT 5 participants, and 12 natives, as depicted in table 4.1:

Table 4.1: Participants and Proficiency

proficiency	number of participants
VLT 2	8
VLT 3	7
VLT 4	6
VLT 5	20
natives	12

4.2 Sentence Analysis Criteria

Since we are dealing with a task in which participants were expected to detect and correct ungrammatical sentences and reproduce grammatical sentences, we decided to set some parameters to analyze the outcome produced by participants. We classified participants' answers as expected when correct, unexpected when incorrect, and NA (not applicable) when answers were either blank or incomplete. Firstly, the sentences are verbal stimuli – that is, participants saw a written sentence for 5,000 msec – and recalled it from its verbal content[58]. Therefore, if they entered a sentence without the verb, we considered it incomplete, namely, NA, as illustrated in (b) and (c) in table 4.2. However, if they stopped on the verb, we recognize it as expected, as in (a). Even though the third scenario below, as represented in (c) in table 4.2, contains most of the sentence elements, it has no verb, namely, the most critical element for our study.

---

[58] We could have a different type of analysis if participants described an image because the stimulus would be non-verbal.

Table 4.2: Expected and NA Answers

Stimulus	Outcome	Classification	Reason
(a) The **boat** for the workers always **charge** a lot of cash	The **boat** for the workers always **charges**	expected	stopped on verb
(b) The **boat** for the workers always **charge** a lot of cash	The **boat** for the workers always	NA	stopped before the verb
(c) The **boat** for the workers always **charge** a lot of cash	The **boat** for the workers always a lot of cash	NA	no verb

Secondly, participants received instruction to correct any mistake they came across. Consequently, ungrammatical sentences corrected to other time tenses were considered expected answers, as exemplified in (d), (e), and (f) in table 4.3. On the other hand, failing to make the sentence grammatical was considered unexpected, as in (h). Lexical changes were also regarded to be expected as long as they were actual words[59], as performed in (e) and (g) but not in (i). Moreover, the typo condition consisted of an inversion of the last two letters of the verb before the morpheme -s. Whenever participants rearranged the letters to form the verb but omitted the morpheme, we considered it to be expected[60], as instantiated in (j). However, the grammatical condition anticipated that participants would correctly reproduce the sentences because they had access to the right input. Should they fail to produce a grammatical sentence, the outcome would be considered unexpected, especially with typos in the verb, as carried out in (k) and (l). Other spelling mistakes did not seem to hinder the main goal of the task, as in (m). Nevertheless, morpheme oversuppliance was considered unexpected, as portrayed in (n) in table 4.3:

[59] Meaning is preserved longer than form (SACHS, 1967). Besides, lexical items and syntactic structures recently activated can prime other ones (LOMBARDI; POTTER, 1992; POTTER; LOMBARDI, 1998).

[60] We decided to do this because, otherwise, we would be adding more constructs to our analysis.

Table 4.3: Expected and Unexpected Answers

Stimulus	**Outcome**	**Classification**	**Reason**
(d) The **boy** with the cameras always **take** pictures	The **boys** with the cameras always **take** pictures	expected	plural agreement
(e) The **machine** for the teachers always **print** many tests	The **machine** always **printed** lots of tests for the teachers	expected	time tense and lexical change
(f) Liam **will to kill** the bird	Liam **wants to kill** the bird.	expected	time tense change
(g) Sam **will to tell** jokes	Sam **will tell** jokes **too**	expected	lexical change
(h) Fred **will to pay** the bill	fred **will pay to** the bill	unexpected	fail to correct
(i) Max always **solevs** the problem	Maxwell always **soluts** the problem	unexpected	invented verb
(j) Max always **solevs** the problem	Max always **solve** the problem	expected	lexical rearranging
(k) Anne always **meets** the brother	Anna always **meet** her brother	unexpected	morpheme omission
(l) Anne always **meets** the brother	Anna always **mets** her brother	unexpected	verb typo
(m) The **girl** with the toys always **wear** beautiful clothes	the **girl** with the toys always **wears** beatiful cloathes	expected	spelling problem not located in the verb
(n) The **video** for the kids always **concern** the parents	The **videos** for the kids always **concerns** to the parents	unexpected	morpheme oversuppliance

4.3 Sentence Results

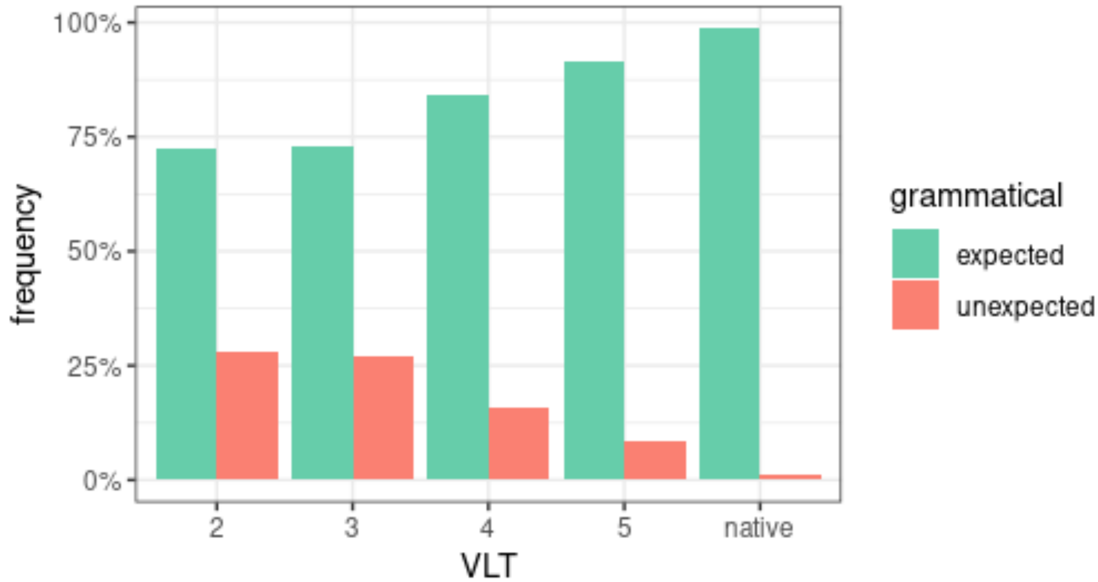
All NA answers were filtered (check Appendix 8 for the analysis script) before the analysis; thus, we are only analyzing expected and unexpected answers. We had initially set the conditions grammatical, typo, and modal as our screening sentences with a threshold of 80% of accuracy; therefore, participants would need to get five out of the six sentences right. However, we found some very problematic results for the lower proficiency group and the natives and decided to delve into these conditions.

We compared the frequency of expected and unexpected answers for each proficiency group for the grammatical, typo, and modal conditions. The frequency of expected answers in the grammatical condition reveals an ascending result as proficiency increases. Level 2 had 72% of expected answers, level 3 73%, level 4 84%, level 5 91%, and natives 99%. We can see in table 4.4 and graph 4.1 that the frequency of expected answers for the VLT 2 and VLT 3 are below the anticipated 80% accuracy. Proficiency helped participants to reproduce grammatical sentences correctly. Thus, lower proficiency levels had the lowest accuracy score when spotting and reproducing the 3rd person singular morpheme.

Table 4.4: Expected and Unexpected Answers for the Grammatical Condition

proficiency	**grammatical answer**	**n (items X participants)**	**frequency**
2	expected	26	0.722
2	unexpected	10	0.278
3	expected	24	0.727
3	unexpected	9	0.273
4	expected	27	0.844
4	unexpected	5	0.156
5	expected	105	0.913
5	unexpected	10	0.0870
native	expected	71	0.986
native	unexpected	1	0.0139

Graph 4.1: Expected and Unexpected Answers for the Grammatical Condition

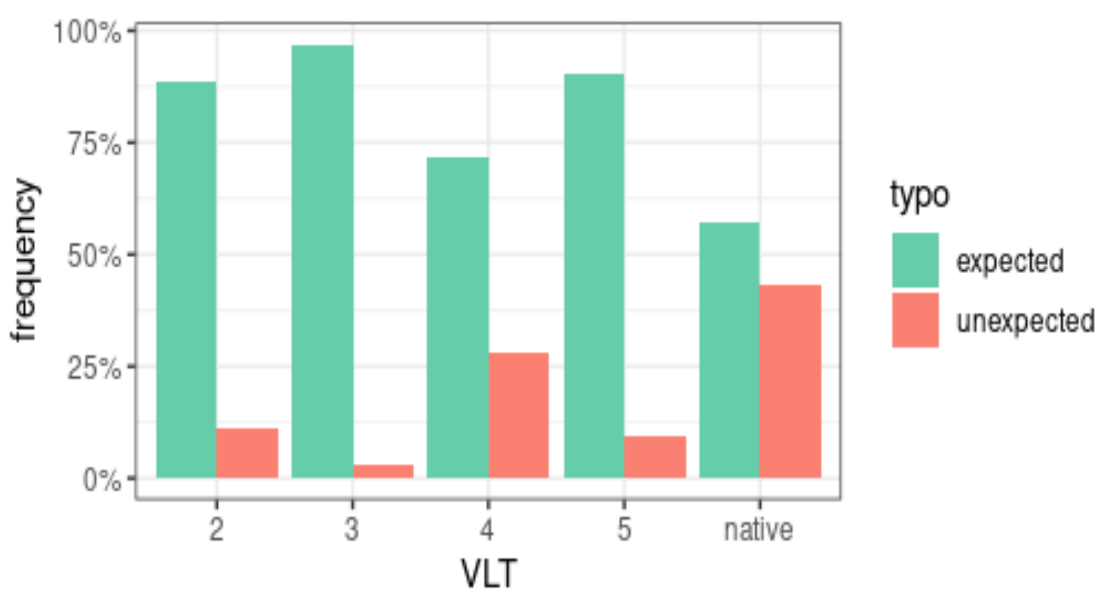


The typo condition showed that the frequency of expected answers was lower than predicted for participants of level 4 and natives, as represented in table 4.5 and graph 4.2. Level 2 participants displayed 89% of expected answers, level 3 97%, level 4 72%, level 5 90%, and the natives 57%. The natives had the lowest rate of expected answers followed by level 4. In this condition, we were led to believe that other elements were influencing the results apart from the proficiency, such as the participants' task understanding.

Table 4.5: Expected and Unexpected Answers for the Typo Condition

proficiency	typo answer	n (items X participants)	frequency
2	expected	32	0.889
2	unexpected	4	0.111
3	expected	32	0.970
3	unexpected	1	0.0303
4	expected	23	0.719
4	unexpected	9	0.281
5	expected	104	0.904
5	unexpected	11	0.0957
native	expected	41	0.569
native	unexpected	31	0.431

Graph 4.2: Expected and Unexpected Answers for the Typo Condition

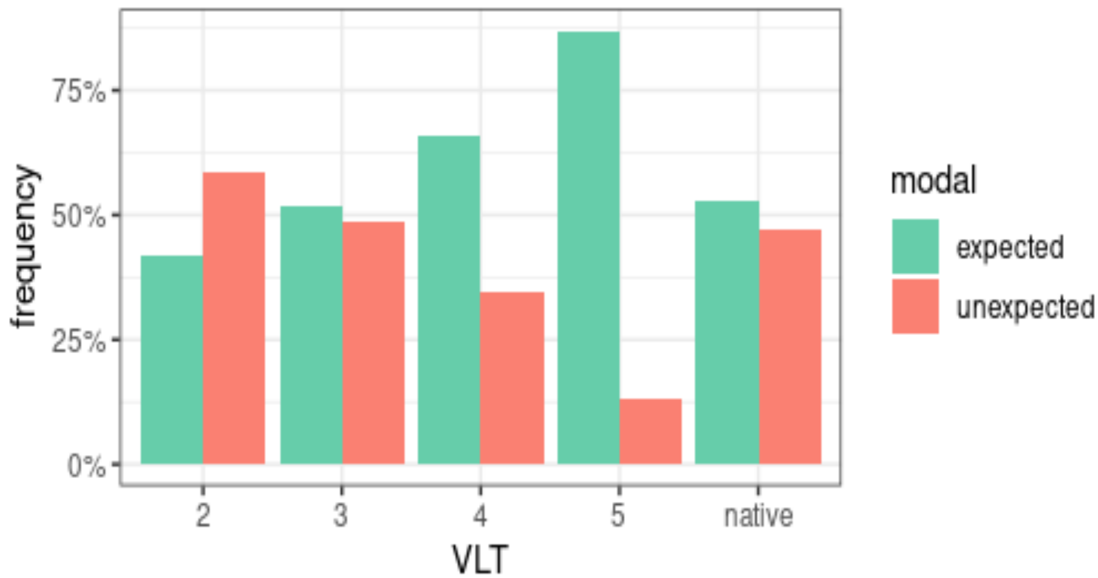


The frequency of expected answers decreased more when we analyzed the modal condition. Expected answers appeared 42% of the time in level 2, 52% in level 3, 66% in level 4, 87% in level 5, and 53% in natives. Only the majority of VLT 5 participants reached the threshold of 80% accuracy, as portrayed in table 4.6 and graph 4.3. Therefore, we had to decide how to deal with these data.

Table 4.6: Expected and Unexpected Answers for the Modal Condition

proficiency	modal answer	n (items X participants)	frequency
2	expected	15	0.417
2	unexpected	21	0.583
3	expected	17	0.515
3	unexpected	16	0.485
4	expected	21	0.656
4	unexpected	11	0.344
5	expected	100	0.870
5	unexpected	15	0.130
native	expected	38	0.528
native	unexpected	34	0.472

Graph 4.3: Expected and Unexpected Answers for the Modal Condition



We decided to investigate if participants had not understood the task demands. After careful consideration and based on some participants' feedback (check appendix 7), we concluded that most participants understood they were expected to correct ungrammatical sentences. As a result, we decided to include all these participants in the study because we had evidence that most comprehended the task, but some groups' overall knowledge enabled them to have a better performance. Besides, we could see that the participants of VLT 2, VLT 3, and VLT 4 displayed similar results; therefore, we decided to group them into lower proficiency. Hence, we have a new configuration of the proficiency groups, as observable in table 4.7 below:

Table 4.7: New Proficiency Configuration

proficiency	**number of participants**
lower	20
higher	21
natives	12

4.4 Inter-rater reliability check

Since we faced a lot of variability in the participants' answers, it was advised that we should have raters rank the sentences produced by participants[61]. We selected three raters for the answers entered by participants in target conditions, animate, inanimate, name, and pronoun, and the control grammatical condition. The raters had to evaluate whether the participants produced an expected or an unexpected answer. No rater had access to the others' answers during the rating process. We examined the raters' converging outcome, meaning that two out of the three had to agree on an expected or unexpected answer, and did the analysis based on it. This helped to solve the variability found in participants' answers. The raters were Brazilian Portuguese-English learners with a background in psycholinguistic studies.

4.5 Target and Control Sentence Conditions Results

In this part, we focus the analysis on our control grammatical condition and the target ungrammatical conditions. Firstly, we present results for each proficiency group for the control and target conditions and then contrast the results between the participants' groups. The results within the participants' groups will help us check our first and second hypotheses and the results between the participants' groups our third and fourth hypotheses.

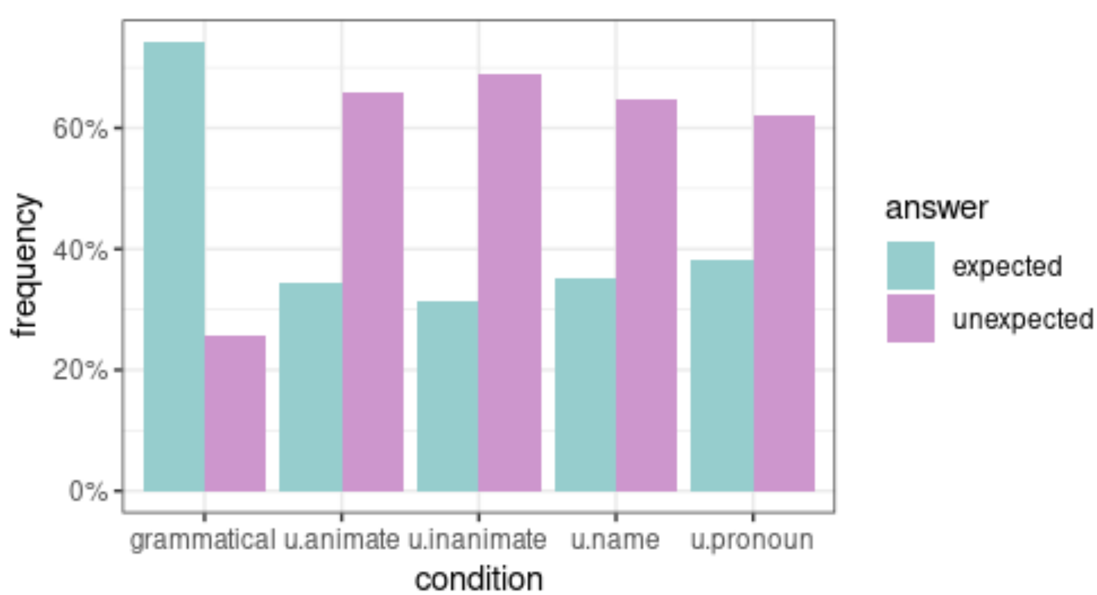
In the lower proficiency group, we can observe, in table 4.8 and graph 4.4, that the frequency of unexpected answers is higher than the expected answers in all ungrammatical conditions. The animate condition has only 34% of expected answers, the inanimate condition 31%, the name condition 35%, and the pronoun 38%. On the other hand, the grammatical condition presents a higher rate of expected answers, 74%, in this group:

[61] This suggestion was proposed by Dr. Mailce Mota.

Table 4.8: Expected and Unexpected Answers for the Lower Proficiency Group

condition	answer	n (items X participants)	frequency
grammatical	expected	92	0.742
grammatical	unexpected	32	0.258
ungrammatical.animate	expected	38	0.342
ungrammatical.animate	unexpected	73	0.658
ungrammatical.inanimate	expected	34	0.312
ungrammatical.inanimate	unexpected	75	0.688
ungrammatical.name	expected	44	0.352
ungrammatical.name	unexpected	81	0.648
ungrammatical.pronoun	expected	48	0.381
ungrammatical.pronoun	unexpected	78	0.619

Graph 4.4: Expected and Unexpected Answers for the Lower Proficiency Group



We adjusted a logistic regression with answer type (expected / unexpected) as response variable and condition (grammatical, ungrammatical.animate, ungrammatical.inanimate, ungrammatical.noun, and ungrammatical.pronoun) as the fixed effect and random intercepts for items and participants. The comparison by nested models revealed that the type of condition contributed significantly to the model ($\chi2 = 110.4$, $p < 0.001$).

We also ran a post-hoc analysis to compare in which conditions the effect was located in the lower proficiency group:

Table 4.9: Contrast between Conditions for the Lower Proficiency Group

contrast	estimate	SE	z.ratio	p.value
grammatical - u.animate	-3.150	0.420	-7.492	<0.0001*
grammatical - u.inanimate	-3.430	0.433	-7.914	<0.0001*
grammatical - u.name	-2.925	0.405	-7.225	<0.0001*
grammatical - u.pronoun	-2.687	0.397	-6.774	<0.0001*
u.animate - u.inanimate	-0.280	0.375	-0.747	0.9453
u.animate - u.name	0.225	0.358	0.627	0.9707
u.animate - u.pronoun	0.463	0.356	1.299	0.6920
u.inanimate - u.name	0.505	0.367	1.374	0.6442
u.inanimate - u.pronoun	0.743	0.366	2.027	0.2528
u.name - u.pronoun	0.238	0.346	0.689	0.9589

* Statistically significant difference

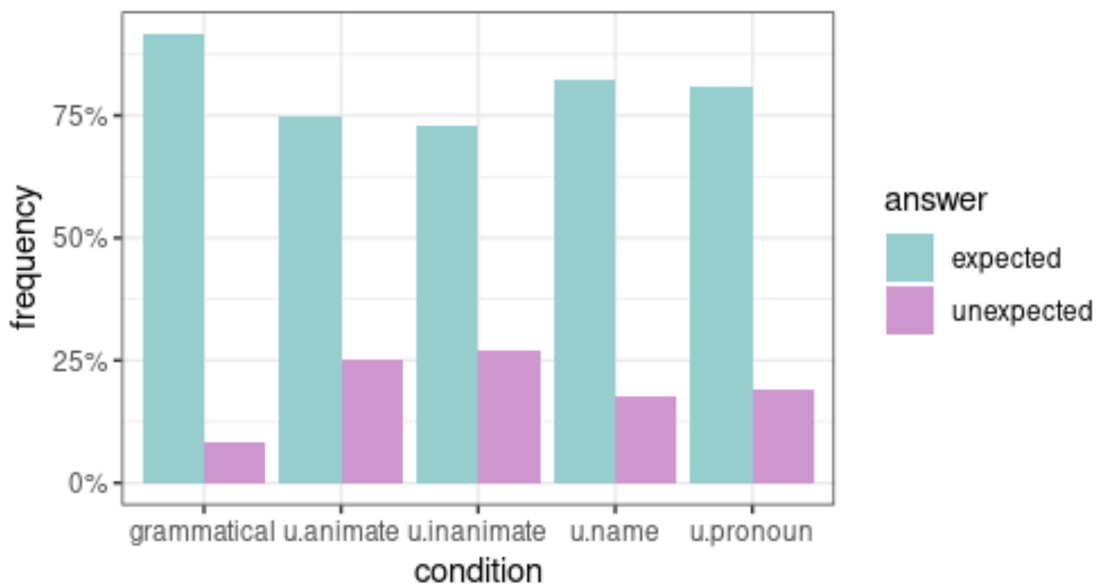
It is possible to see, from table 4.9 above, that the difference was located between grammatical (control) and ungrammatical (target) conditions. There was no significant difference between any ungrammatical conditions for the lower proficiency group. Therefore, our first hypothesis was confirmed for this group since grammatical sentences are easier to reproduce than it is to correct ungrammatical sentences. Erlam (2006) defends that the latter is more challenging, especially in lower proficiency groups. On the other hand, local-distance sentences seemed to be as difficult to correct as the long-distance agreement sentences for this group. Consequently, our second hypothesis was not confirmed because this group had problems with the agreement in local and long-distance sentences. Hoshino, Dussias & Kroll (2010) and Jensen et al. (2019) defend that long-distance agreement is hard to spot and achieve, particularly for lower proficient participants, but the local-distance agreement was also a challenge for this group.

When we checked the frequency of answers in the higher proficiency group, we could see that the rate of expected answers is more frequent than unexpected answers for the target conditions and the control condition, as depicted in table 4.10 and graph 4.5. For the target conditions, the animate condition presents 75% of expected answers, the inanimate condition 73%, the name condition 82%, and the pronoun 81%, and control grammatical condition 92%:

Table 4.10: Expected and Unexpected Answers for the Higher Proficiency Group

condition	**answer**	**n (items X participants)**	**frequency**
grammatical	expected	109	0.916
grammatical	unexpected	10	0.0840
ungrammatical.animate	expected	89	0.748
ungrammatical.animate	unexpected	30	0.252
ungrammatical.inanimate	expected	86	0.729
ungrammatical.inanimate	unexpected	32	0.271
ungrammatical.name	expected	98	0.824
ungrammatical.name	unexpected	21	0.176
ungrammatical.pronoun	expected	97	0.808
ungrammatical.pronoun	unexpected	23	0.192

Graph 4.5: Expected and Unexpected Answers for the Higher Proficiency Group



We did the same adjusted logistic regression with answer type as response variable and condition as the fixed effect and random intercepts for items and participants for the higher proficiency group. The comparison by nested models indicated that the type of condition was substantial to the model ($\chi 2 = 39.58$, $p < 0.001$).

We also conducted a post-hoc analysis to contrast the conditions for the higher proficiency group:

Table 4.11: Contrast between Conditions for the Higher Proficiency Group

contrast	estimate	SE	z.ratio	p.value
grammatical - u.animate	-3.009	0.641	-4.696	<0.0001*
grammatical - u.inanimate	-3.327	0.644	-5.164	<0.0001*
grammatical - u.name	-1.863	0.618	-3.013	0.0218*
grammatical - u.pronoun	-2.035	0.619	-3.286	0.0090*
u.animate - u.inanimate	-0.318	0.474	-0.670	0.9628
u.animate - u.name	1.146	0.547	2.094	0.2224
u.animate - u.pronoun	0.974	0.537	1.815	0.3645
u.inanimate - u.name	1.464	0.543	2.695	0.0547
u.inanimate - u.pronoun	1.292	0.531	2.432	0.1070
u.name - u.pronoun	-0.172	0.558	-0.308	0.9980

* Statistically significant difference

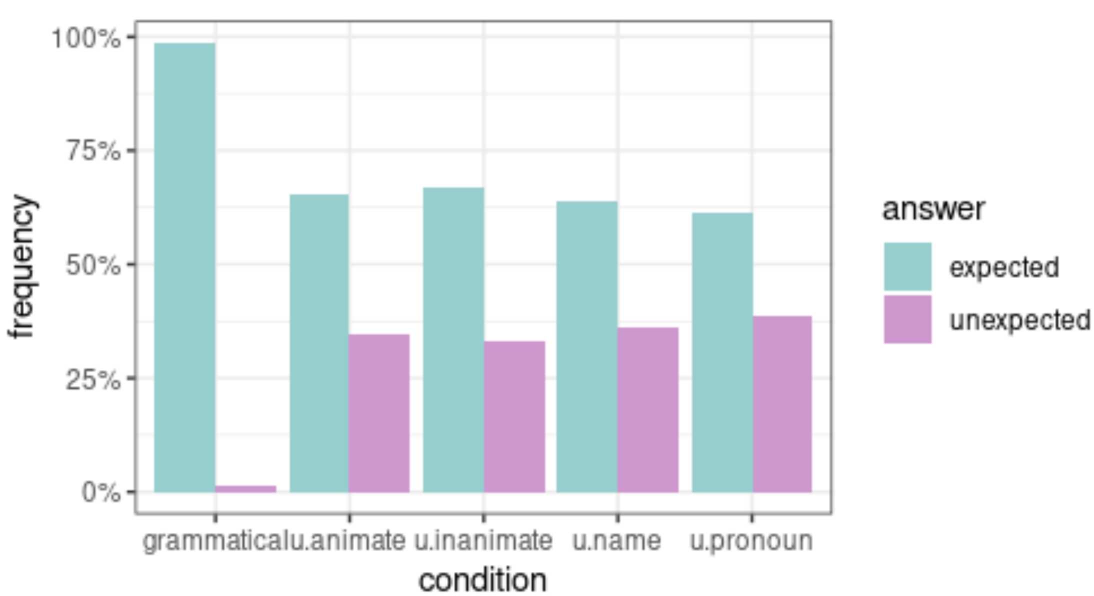
Such as we observed in the lower proficiency group, the difference in the higher proficiency group, as illustrated in table 4.11 above, was detected between the grammatical condition and the ungrammatical conditions. Our first hypothesis was also corroborated for this group because the grammatical condition presented a higher rate of expected answers than the ungrammatical conditions. However, this group had a fair amount of expected answers in both local-distance and long-distance ungrammatical sentences. Consequently, our second hypothesis was rejected because this group had no problem with different types of NPs distance.

Contrary to what we expected, the native group displayed a lower rate of expected answers in the target conditions in relation to the higher proficiency group, as we can check in table 4.12 and graph 4.6. This happened for the target conditions as the animate condition displays 65% of expected answers, the inanimate condition 67%, the name condition 64%, and the pronoun 61%, but for the control grammatical condition exhibits 99% of expected answers:

Table 4.12: Expected and Unexpected Answers for the Native Group

condition	answer	n (items X participants)	frequency
grammatical	expected	71	0.986
grammatical	unexpected	1	0.0139
ungrammatical.animate	expected	47	0.653
ungrammatical.animate	unexpected	25	0.347
ungrammatical.inanimate	expected	48	0.667
ungrammatical.inanimate	unexpected	24	0.333
ungrammatical.name	expected	46	0.639
ungrammatical.name	unexpected	26	0.361
ungrammatical.pronoun	expected	44	0.611
ungrammatical.pronoun	unexpected	28	0.389

Graph 4.6: Expected and Unexpected Answers for the Native Proficiency Group



We carried out the same kind of adjusted logistic regression with answer type as response variable and condition as the fixed effect and random intercepts for items and participants for the native group. The nested models' comparison showed that the type of condition was crucial to the model ($\chi2 = 95.138$, $p < 0.001$).

We performed a post-hoc analysis to contrast the conditions for the native group:

Table 4.13: Contrast between Conditions for the Native group

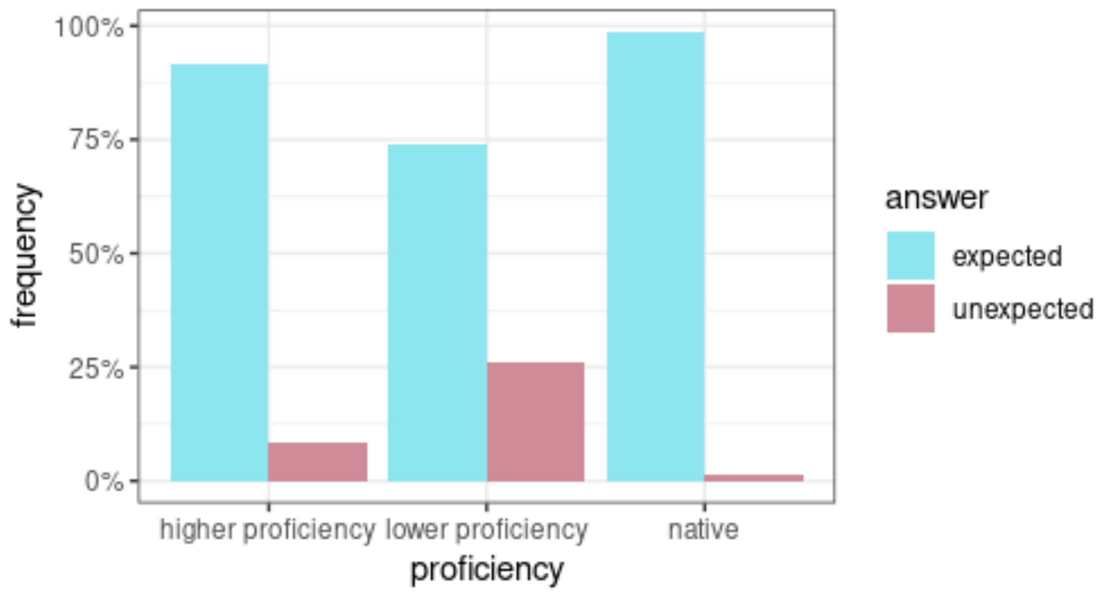
contrast	estimate	SE	z.ratio	p.value
grammatical - u.animate	-8.235	2.218	-3.712	0.0019*
grammatical - u.inanimate	-7.980	2.206	-3.618	0.0027*
grammatical - u.name	-8.492	2.231	-3.807	0.0013*
grammatical - u.pronoun	-9.012	2.254	-3.997	0.0006*
u.animate - u.inanimate	0.254	0.714	0.356	0.9966
u.animate - u.name	-0.258	0.719	-0.358	0.9965
u.animate - u.pronoun	-0.777	0.730	-1.065	0.8246
u.inanimate - u.name	-0.512	0.720	-0.711	0.9540
u.inanimate - u.pronoun	-1.032	0.735	-1.403	0.6260
u.name - u.pronoun	-0.520	0.725	-0.716	0.9528

* Statistically significant difference

As previously observed for the bilingual groups, only the difference between the grammatical condition and the ungrammatical conditions was significant for the native group, as represented in table 2.13 above. Hence, our first hypothesis was also proved for this group. However, our second hypothesis was not confirmed because the rate of expected answers in local-distance and long-distance conditions were very similar.

Afterwards, we start comparing the results between the groups for each condition. We begin with our control grammatical condition that will help us test our third hypothesis and check whether there is a greater demand for WMC for the lower proficiency group when they reproduce grammatical input. In graph 4.7, we see the frequency of expected and unexpected answers in this condition for each group. Natives have the highest score followed by the higher proficiency group and then the lower proficiency group:

Graph 4.7: Expected and Unexpected Answers in Grammatical Condition



We adjusted a logistic regression for the grammatical condition with answer type (expected / unexpected) as response variable and proficiency (higher proficiency, lower proficiency, and native) as the fixed effect and random intercepts for items and participants. The comparison by nested models demonstrated that the groups' proficiency contributed to the model ($\chi2 = 12.529$, $p = 0.001902$).

We performed a post-hoc analysis, as exemplified in table 4.14, to contrast the performance between groups for the grammatical condition

Table 4.14: Contrast between the Groups in the Grammatical Condition

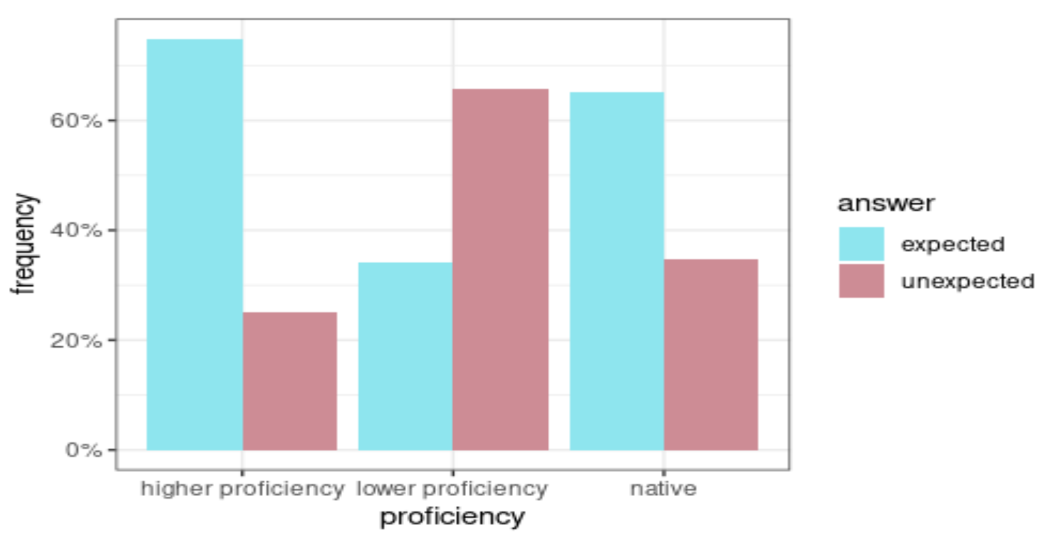
contrast	estimate	SE	z.ratio	p.value
higher proficiency - lower proficiency	-2.38	1.03	-2.317	0.0535
higher proficiency - native	1.74	1.53	1.142	0.4882
lower proficiency - native	4.12	1.52	2.718	0.0181*

* Statistically significant difference

We found a significant difference between the performance of the native and lower proficiency groups (z= 2.718, p= 0.0181) and a marginal difference between the performance of the higher proficiency and lower proficiency groups (z= -2.317, p= 0.0535) in the grammatical condition. Therefore, our third hypothesis was partially confirmed because there was only a marginal difference between the performance of higher and lower proficiency groups but a significant difference between the performance of native and lower proficiency groups. WMC may have been an issue that facilitated the performance of the groups in reproducing grammatical sentences because they needed to perceive, store, and reproduce the 3rd person singular morpheme (-s). Nonetheless, participants also had to use their knowledge in English. WMC predicted the listening comprehension performance of participants in Azevedo's (2012) study and it also enabled retention and acquisition of a complex syntactic form in Finardi's (2009) study. Furthermore, WMC is a key element in SLA (WEN; MOTA; MCNEILL, 2013).

Following the analysis of the target ungrammatical conditions, we will be able to check our fourth hypothesis and discover if proficiency influences the participants' ability to detect and correct ungrammatical sentences. We will need to go over our four target conditions to have an answer for this hypothesis. We conducted the same type of analysis, as in the grammatical condition, for the target conditions. We will start by analyzing the animate condition, which is one of our long-distance agreement conditions. In graph 4.8, we can observe the performance of the three groups in the animate condition. The higher proficiency group has the highest rate of expected answers, the natives in second and the lower proficiency group in last:

Graph 4.8: Expected and Unexpected Answers in Animate Condition



A logistic regression was adjusted for the animate condition with answer type as response variable and proficiency as the fixed effect and random intercepts for items and participants. The nested models' comparison confirmed that the levels of proficiency contributed to the model ($\chi 2$ = 12.815, p = 0.001649).

We also did a post-hoc analysis to compare the performance between groups for the animate condition, as illustrated in table 4.15 below:

Table 4.15: Contrast between the Groups in the Animate Condition

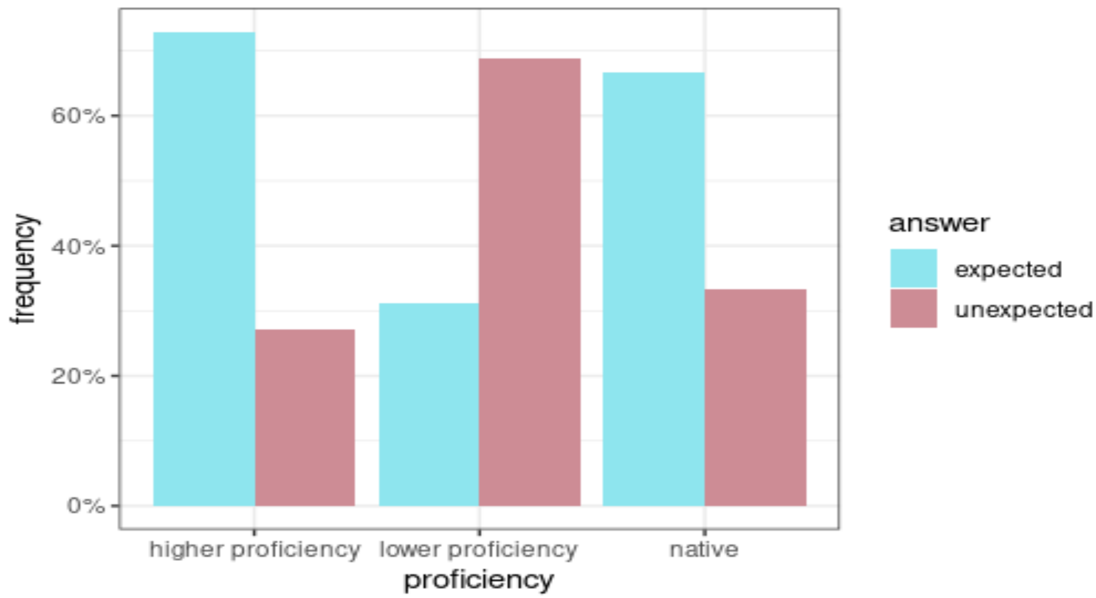
contrast	estimate	SE	z.ratio	p.value
higher proficiency - lower proficiency	-4.290	1.41	-2.317	0.0066*
higher proficiency - native	-0.799	1.49	-0.537	0.8531
lower proficiency - native	3.491	1.58	2.205	0.0703

* Statistically significant difference

Only the difference between the performance of the higher proficiency and lower proficiency groups was significant (z= -2.317, p= 0.0066), and not even the difference between the native and lower proficiency groups was statistically significant (z= 2.205, p= 0.0703) in the animate condition.

We present the results for our second long-distance agreement condition, namely inanimate condition. We can examine the groups' performance in the inanimate condition in graph 4.9. It has a similar pattern observed in the animate condition with the higher proficiency being on the top, the natives in the middle, and the lower proficiency group on the bottom:

Graph 4.9: Expected and Unexpected Answers in Inanimate Condition



An adjusted logistic regression for the inanimate condition was done with answer type as response variable and proficiency as the fixed effect and random intercepts for items and participants. The comparison by nested models also verified that the levels of proficiency contributed to the model ($\chi2 = 15.831$, $p = 0.0003651$).

A post-hoc analysis (table 4.16) to compare the performance between groups for the inanimate condition was carried out:

Table 4.16: Contrast between the Groups in the Inanimate Condition

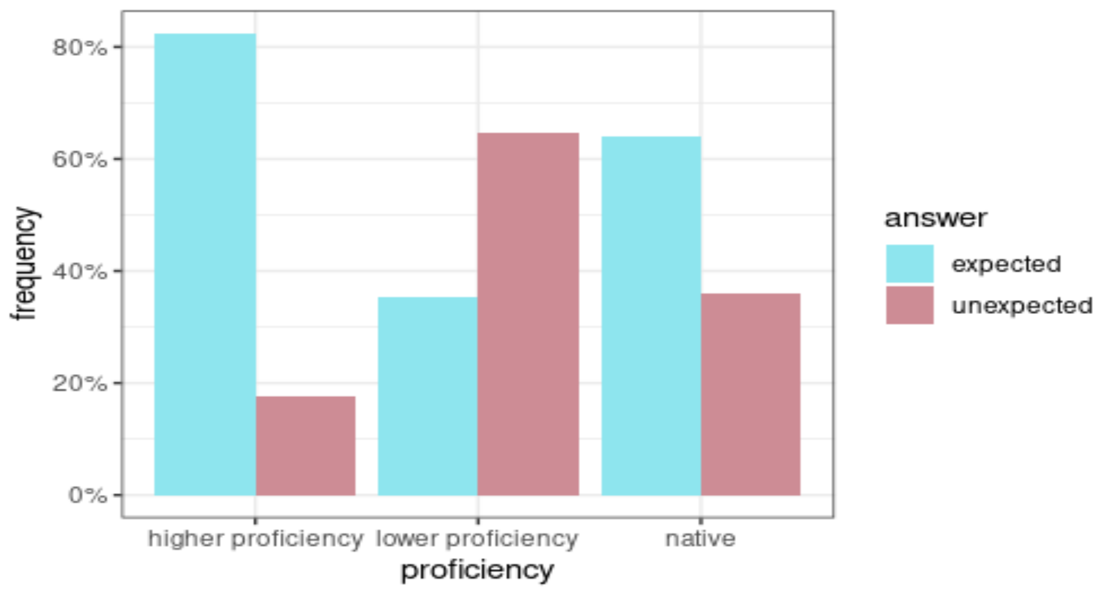
contrast	estimate	SE	z.ratio	p.value
higher proficiency - lower proficiency	-3.418	0.973	-3.514	0.0013*
higher proficiency - native	-0.522	1.044	-0.500	0.8712
lower proficiency - native	2.896	1.079	2.684	0.0199*

* Statistically significant difference

A significant difference was found in this condition between the performance of the higher proficiency and lower proficiency groups (z=-3.514, p= 0.0013) and between the performance of the native and lower proficiency groups (z=2.684, p= 0.0199).

We reached the analysis for our first local-distance agreement condition, the name condition. We can inspect the outcomes produced by the three groups in the name condition in graph 4.10. Once more, the higher proficiency group ranks the best performance, the natives second, and the lower proficiency group third:

Graph 4.10: Expected and Unexpected Answers in Name Condition



We adjusted a logistic regression for the name condition with answer type as response variable and proficiency as the fixed effect and random intercepts for items and participants. The nested models' comparison demonstrated that the difference displayed in each proficiency group contributed to the model ($\chi 2 = 20.726$, $p < 0.001$).

We compared the performance between groups for the name condition through a post-hoc analysis, the results are shown in table 4.17 below:

Table 4.17: Contrast between the groups in the Name Condition

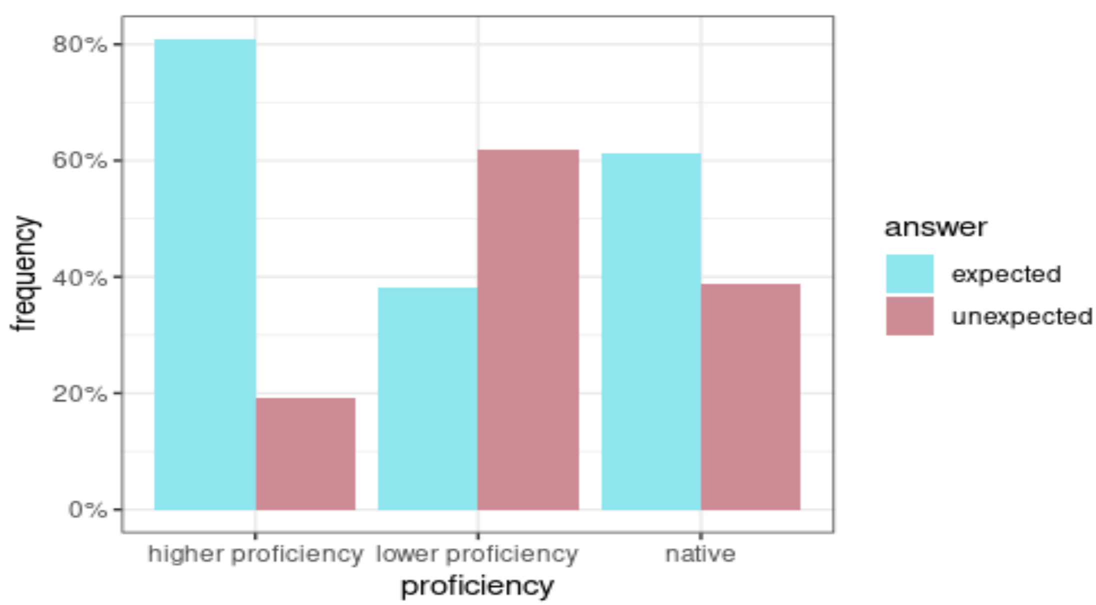
contrast	estimate	SE	z.ratio	p.value
higher proficiency - lower proficiency	-8.62	2.60	-3.321	0.0026*
higher proficiency - native	-2.85	2.45	-1.160	0.4775
lower proficiency - native	5.78	3.51	1.644	0.2272

* Statistically significant difference

Only the difference between the performance of the higher proficiency and lower proficiency groups was significant (z= -3.321, p= 0.0026) in the name condition.

Finally, we present the analysis for the pronoun condition, our second local-distance agreement condition. Graph 4.11 describes the results of the three groups in the pronoun condition. The higher proficiency group performed best, the natives' group follows right after, and at last, we have the lower proficiency group:

Graph 4.11: Expected and Unexpected Answers in Pronoun Condition



An adjusted logistic regression for the pronoun condition was conducted with answer type as response variable and proficiency as the fixed effect and random intercepts for items and participants. The comparison by nested models suggested that the difference in the proficiency groups contributed to the model ($\chi2$ =22.626, p <0.001).

We used a post-hoc analysis to contrast the performance between groups for the pronoun condition, as displayed in table 4.18:

Table 4.18: Contrast between the Groups in the Pronoun Condition

contrast	estimate	SE	z.ratio	p.value
higher proficiency - lower proficiency	-15.90	3.41	-4.664	<0.0001*
higher proficiency - native	-1.36	2.33	-0.584	0.8285
lower proficiency - native	14.54	3.77	3.853	0.0003*

* Statistically significant difference

The difference between the performance of the higher proficiency and lower proficiency groups (z= -4.664, p <.0001) and the performance of native and lower proficiency groups (z= 3.853, p = 0.0003) were significant in the pronoun condition.

As a result, our fourth hypothesis was partially confirmed. While the higher proficiency group outperformed the lower proficiency participants in detecting and correcting all the ungrammatical sentences, the natives did not perform significantly better than the lower proficiency group in the animate and the name conditions. The bilinguals' performance was mediated by proficiency (ERLAM, 2006), but the natives' performance exhibited optionality, which is found in some English varieties (LABOV et al., 1968). It seems that only the L2 learners that had advanced L2 lexical knowledge could access and assemble this morphological information (JIANG, 2000; JACKENDOFF & AUDRING, 2016). Proficiency also helped predict the L2 performance in detecting and correcting ungrammatical sentences (HOSHINO; DUSSIAS; KROLL, 2010; JENSEN et al., 2019).

Furthermore, we adjusted a logistic regression for answer type as response variable and the interaction between proficiency and type of condition, and random intercepts for items and participants. The nested models' comparison revealed that the interaction between proficiency and condition was significant to the model ($\chi2$ = 25.669, p = 0.001197).

We ran a post-hoc analysis for this interaction between conditions and proficiency groups (check Appendix 9). This type of analysis accounted for the comparison between groups and conditions including different groups and different conditions. Besides some significant differences between grammatical conditions in a group and a target condition in others, we found some statistically significant results, such as the difference between the higher proficiency animate condition and lower proficiency inanimate condition ($z= -3.499$, $p= 0.0352$), the lower proficiency animate condition and the higher proficiency name condition ($z= 4.213$, $p= 0.0023$), the lower proficiency inanimate condition and higher proficiency name condition ($z= 4.466$, $p= 0.0008$), the lower proficiency inanimate condition and the higher proficiency pronoun condition ($z= 4.327$, $p= 0.0014$), the higher proficiency name condition and lower proficiency name condition ($z= -4.012$, $p= 0.0054$), the higher proficiency name condition and lower proficiency pronoun condition ($z= -3.789$, $p= 0.0127$), the lower proficiency name condition and higher proficiency pronoun condition ($z= 3.870$, $p= 0.0094$), the higher proficiency pronoun condition and lower proficiency pronoun condition ($z= -3.645$, $p= 0.0214$). Our main focus was the contrast within groups considering the control and the target conditions and between groups considering each condition. Nevertheless, it is important to consider that there are differences between different groups considering different conditions.

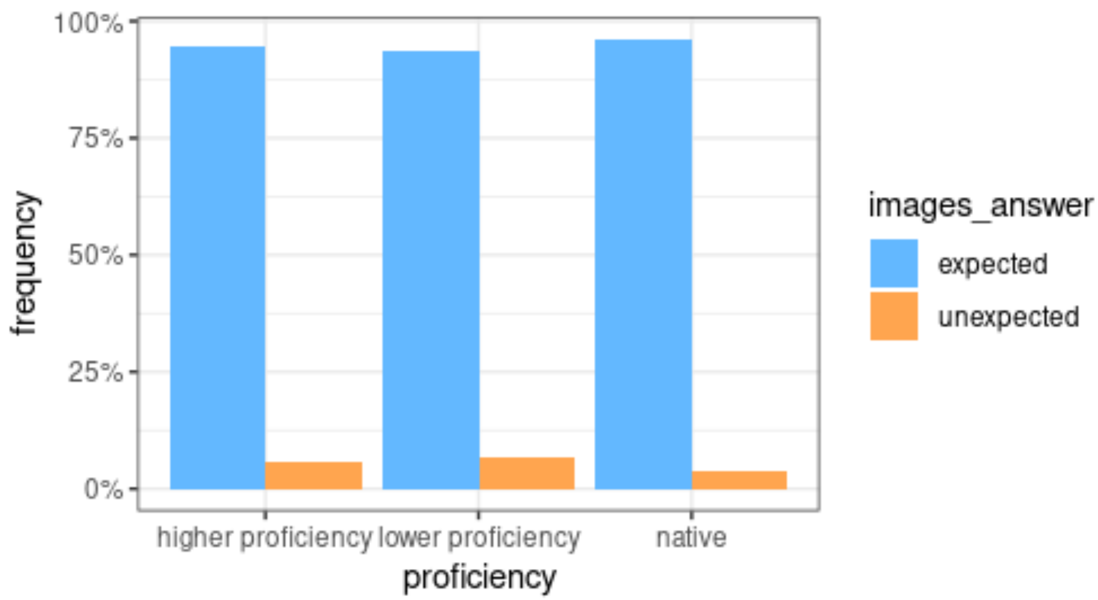
4.6 Images and Letters Results

We grouped the outcome of the sequences of images and letters participants judged into expected when they got the sequence right, unexpected when they got it wrong, and NA whenever they missed the sequence because of the time constraint of 10,000 msec. We also filtered NA answers (check Appendix 8) and conducted the analysis with the expected and unexpected answers of the higher proficiency, lower proficiency, and native groups. The images and letters were all classified into images. This analysis will help us test our fifth hypothesis and discover if proficiency impacted the groups' performance on image recall and judgment. Table 4.19 and graph 4.12 illustrate the expected and unexpected answers that the groups produced in judging the sequences of images and letters. The natives had 96% of expected answers, the higher proficiency group 95%, and the lower proficiency group 93%:

Table 4.19: Expected and Unexpected Answers for the Images and Letters

group	images answer	n (items X participants)	frequency
higher proficiency	expected	790	0.945
higher proficiency	unexpected	46	0.0550
lower proficiency	expected	819	0.934
lower proficiency	unexpected	58	0.0661
native	expected	484	0.960
native	unexpected	20	0.0397

Graph 4.12: Expected and Unexpected Answers for the Images and Letters



As done for the sentences, we adjusted a logistic regression for the images and letters with answer type (expected / unexpected) as response variable and proficiency (higher proficiency, lower proficiency, and native) as the fixed effect and random intercepts for items and participants. The comparison by nested models pointed out that proficiency contributed to the model ($\chi2$ =31.75, p <0.001).

We ran a post-hoc analysis to compare the performance between the groups in judging the images and letters, as presented in table 4.20:

Table 4.20: Contrast between the Groups for the Images

contrast	estimate	SE	z.ratio	p.value
higher proficiency - lower proficiency	-47.2	5.43	-8.690	<0.0001
higher proficiency - native	20.9	6.65	3.135	0.0049
lower proficiency - native	68.0	9.85	6.906	<0.0001

* Statistically significant difference

The contrast between the groups was statistically significant. The higher proficiency group performed significantly better than the lower proficiency group ($z=8.690$, $p<0.0001$), the native performed significantly differently from the higher proficiency group ($z= 3.135$, $p= 0.0049$), and the lower proficiency group ($z= 6.906$, $p<0.0001$).

The fifth hypothesis was not confirmed because the contrast between the performance of the groups was statistically significant. WMC is used in rehearsal, noticing, attention regulation, and task-switching (WEN; MOTA; MCNEILL, 2013). Learners' WMC may have impacted the linguistic cues on which they focused their attention. Switching between sentence and image recall could have added an extra load to less proficient participants and impaired their ability to perceive and recall the 3rd person singular morpheme (-s) in grammatical sentences and detect and correct its omission in ungrammatical sentences. WMC could have fulfilled a function in the L2 groups' performance (FINARDI, 2009; AZEVEDO, 2012).

4.7 Summary of Results

In table 4.21, we present a summary of our hypotheses and results:

Table 4.21: Summary of Hypotheses and Results

Hypotheses	**Results**	**Confirmed/ Rejected**
a. Grammatical sentences with 3rd person singular morpheme are more accurately reproduced than correcting ungrammatical sentences without the 3rd person singular morpheme within the groups.	All three groups performed better at reproducing grammatical sentences than correcting ungrammatical sentences.	confirmed
b. Long-distance agreement ungrammatical sentences without the 3rd person singular morpheme are harder to detect and correct than the local-distance agreement ungrammatical sentences without the 3rd person singular morpheme within the groups.	The higher proficiency group performed similarly well, 73-82%, in the ungrammatical conditions, the lower proficiency group displayed performance of 31-38% in the ungrammatical conditions, and natives 61-65% in these conditions.	rejected
c. There is greater demand for working memory capacity for the lower proficiency group, affecting the reproduction of grammatical sentences with the 3rd person singular morpheme compared to the other groups.	The difference between the native group and the lower proficiency group was statistically significant in reproducing grammatical sentences, but the difference between the higher proficiency and the lower proficiency groups was only marginally different.	partially confirmed
d. The proficiency level influences the participants' ability to detect and correct ungrammatical sentences without the 3rd person singular morpheme between groups.	While higher proficiency participants outperformed lower proficient participants in all ungrammatical conditions, the natives were not statistically better than the lower proficiency group in the animate and name conditions.	partially confirmed
e. Proficiency does not impact image recall and judgment between the groups.	The contrast between the groups was statistically different in image recall and judgment.	rejected

CHAPTER 5

DISCUSSION

The goal of this study was to check the lower proficiency, higher proficiency, and native groups' performance in a task in which they had to reproduce grammatical sentences and detect and correct ungrammatical sentences with the expression of 3rd person singular morpheme (-s). We distributed our target conditions into ungrammatical sentences with long-distance agreement – animate and inanimate – and local-distance agreement – name and pronoun – and investigated whether proficiency influenced the results. Moreover, we examined if the WMC of these groups helped them reproduce grammatical sentences accurately and judge sequences of images. Thus, we focused on (i) participants' ability to retrieve, manipulate, and produce 3rd person singular morpheme and (ii) participants' performance in a linguistic and non-linguistic WM task. We discuss our findings based on the theoretical background presented in this study.

Our first hypothesis, which states that grammatical sentences with 3rd person singular morpheme are more accurately reproduced than correcting ungrammatical sentences without the 3rd person singular morpheme within the groups, was confirmed. The three groups had a higher rate of expected answers than unexpected answers in the grammatical condition (the lower proficiency group had 74% of expected answers, the higher proficiency group 92%, and the native group 99%). Hence, participants displayed a fair amount of retention and reproduction of the 3rd person singular morpheme (-s). According to Erlam (2006), correcting ungrammatical sentences is more troublesome than reproducing grammatical sentences. Indeed, that is what we observed, as this scenario of a high rate of verbatim recollections of the grammatical construction changed a lot when we checked the ungrammatical conditions for some groups.

We could not confirm our second hypothesis: long-distance agreement ungrammatical sentences without the 3rd person singular morpheme are harder to detect and correct than the local-distance agreement ungrammatical sentences without the 3rd person singular morpheme within the groups. The results within the groups display a similar rate of expected and unexpected answers in all the ungrammatical conditions with long-distance and local-distance agreements. In Erlam's (2006) study, natives accurately reproduced grammatical sentences 97% of the time, and 91% of ungrammatical sentences were corrected. Her L2 learners managed to repeat 61% of grammatical sentences and correct 35% of ungrammatical sentences. Besides giving explicit instruction for participants to correct the sentences, Erlam (2006) also provided feedback on participants' answers during training. Some participants from our native group may have failed to understand the task because their rate of ungrammatical sentence correction varied from 61% in the pronoun condition, 64% in the name condition, 65% in the animate condition to up to 67% in the inanimate condition. The lower proficiency group was able to correct 31% of the inanimate condition, 34% of the animate condition, 35% of the name condition, and 38% in the pronoun condition. Yet, our higher proficiency group corrected 75% of the animate condition, 73% of the inanimate condition, 82% of the name condition, and 81% of the pronoun condition. The higher proficiency group displayed the best performance in the target conditions, but long-distance agreement sentences are very troublesome for low-proficient learners (HOSHINO; DUSSIAS; KROLL, 2010; JENSEN et al., 2019). In our study, the lower proficient group had problems with all ungrammatical conditions independently of the agreement distance.

Our third hypothesis, there is greater demand for working memory capacity for the lower proficiency group, affecting the reproduction of grammatical sentences with the 3rd person singular morpheme compared to the other groups., was partially confirmed. However, the fifth hypothesis, proficiency does not impact image recall and judgment between the groups, was not corroborated because the non-linguistic task was an issue dictated by proficiency. The performance of the higher proficiency and the lower proficiency groups in the grammatical condition was marginally different ($z= -2.317$, $p= 0.0535$), and the difference between the performance of the native and the lower proficiency groups was statistically significant ($z= 2.718$, $p= 0.0181$). Nonetheless, in the image recall and judgment, contrasts between the higher proficiency and lower proficiency groups($z= -8.690$, $p<0.0001$), between the native and the higher proficiency groups ($z= 3.135$, $p=0.0049$), and between the native and the lower proficiency groups ($z= 6.906$, $p<0.0001$) were statistically significant. WM is important for performing linguistic tasks and is used in task switching. WMC can influence L2 learners in tasks that demand their L2 language knowledge (WEN; MOTA; MCNEILL, 2013), and it can help predict L2 learners' performance in linguistic tasks (FINARDI, 2009; AZEVEDO, 2012), this can also be true for anticipating participants' ability to reproduce grammatical sentences accurately. A stimulus can trigger items that are stored in LTM, but if the knowledge is not internalized yet, it cannot be retrieved. Awareness can be used for explicit recall; however, no information can be maintained in the focus of attention if it is not activated. WM is supported by activated items from LTM; therefore, information processing is dependent on information stored in LTM (COWAN, 1988). Encoding and retrieval resources are very essential for accurate performances (WEN; MOTA; MCNEILL, 2013). The results indicate that information about 3rd person singular morpheme production was not ready to be accessed by our lower proficiency group (JIANG, 2000) when they had to supply the morpheme omission. Even though lower proficiency participants managed to recall the ungrammatical sentences, they did not have enough knowledge to correct them (SCHIMITT, 2014).

We partially accept the fourth hypothesis, which argues that the proficiency level influences the participants' ability to detect and correct ungrammatical sentences without the 3rd person singular morpheme between groups. The higher proficiency group outperformed lower proficiency in our target conditions with long-distance – animate ($z= -2.317$, $p= 0.0066$) and inanimate ($z=-3.514$, $p= 0.0013$) – and local-distance – name ($z=-3.321$, $p= 0.0026$) and pronoun ($z=-4.664$, $p= <0.0001$) agreement sentences. Thus, correcting the sentences was mediated by proficiency independently of their type (ERLAM, 2006). Higher proficiency L2 learners may have reached a higher level of lexical development; thus, they were able to retrieve and encode morphological information to correct the ungrammatical sentences (JIANG, 2000; JACKENDOFF & AUDRING, 2016). Differently from what we anticipated, the natives did not perform significantly differently than the lower proficient group in the animate condition ($z= 2.205$, $p=0.0703$) and the name condition ($z= 1.644$, $p= 0.2272$). In some English varieties, people choose to omit the 3rd person singular agreement mark (LABOV et al., 1968). This may help explain why natives had no statistically significant better results than the lower proficiency participants in these two conditions. Besides, animacy did not seem to be an important factor in helping participants correct the sentences because the animate condition did not yield substantially different results from the inanimate condition between natives and bilingual participants. The difference in name and pronoun conditions was not significant between any group either.

Our data enables us to understand more about L2 learners' ability to detect and correct functional morphology violations and reproduce sentences with functional morphology. Slabakova (2013, 2014) defends that functional morphology is a challenge for L2 learners in production and comprehension as it is the bottleneck of language acquisition. Processing difficulties in functional morphology have been found in L2 learners from different native languages and different proficiency levels (cf. JIANG, 2004, 2007; CARNEIRO, 2008, 2011; HOSHINO; DUSSIAS; KROLL, 2010; MACWHINNEY, 2017; JENSEN et al., 2019; OLIVEIRA; FONTOURA &

SOUZA, 2020). This type of problem seems to increase in processing long-distance agreement sentences (JENSEN et al., 2019). Proficiency is a major factor when L2 learners have to produce sentences with conceptual agreement (HOSHINO; DUSSIAS; KROLL, 2010). The results from our experiment indicate that proficiency was essential to enable participants to correct ungrammatical sentences because the higher proficiency group was significantly better than lower proficiency in all four target conditions. However, the higher proficiency group was only marginally better than the lower proficiency group in reproducing the grammatical sentences.

The results seem to be in accordance with a shared network of processing and representation of morphemes (JACKENDOFF & AUDRING, 2016). Learners that have no internalized knowledge of morphemes do not seem to have this shared network integrated. If one cannot process and retain this subtle information in the morphemes, the person will not be able to produce them. The entrenchment faced by L2 learners (MACWHINNEY, 2017) can only be overcome at higher levels of proficiency; therefore, it is only at the end of the stages of lexical development that an L2 learner has access to all lemma and lexeme information (JIANG, 2000). Adult L2 learners may have problems with such analytical forms in the morphemes (MACWHINNEY, 2017).

Monitoring one's production is not an easy task (KRASHEN, [1981] 2002). One can have time to plan the written message, but if they have not internalized this type of knowledge, they will not be able to produce grammatical agreement sentences. Learners focus more on the message rather than on formal aspects; thus, morphological aspects seem to be left aside because they have a very restricted role in performance. In natural contexts, such as in immersed bilinguals, inferencing may happen for learning morphological cues (CARTON, 1971), but non-immersed bilinguals may not have enough meaningful exposure to learning these subtle cues. Despite being introduced at an early stage of formal instruction, the 3rd person singular morpheme is very troublesome for L2 learners. Instructors may need to do more than reinforce it during classes because the learnability of a structure can be influenced by the amount of exposition. As a result, frequency of exposition is essential (LARSEN-FREEMAN, 1976). Jiang (2000) also enforces that lexical development needs enough exposure, and morphological encoding is the last stage of lexical development.

Perceiving the 3rd person singular morpheme presence, storing it, and reproducing it demands WM mechanisms that need to be retrieved from LTM. WM assembles and processes information stored in LTM. Spreading activation can decrease the WM load by the linkage between items (JACKENDOFF & AUDRING, 2016). Should L2 knowledge have not been fully developed yet, L2 learners fail to reproduce grammatical input in the form of morphological specifications, as we observed for some lower proficiency participants (JIANG, 2000). WMC may direct learners' attention to linguistic cues during the learning process. This could be a meaningful resource in linguistic tasks (WEN; MOTA, MCNEILL, 2013). Besides, the non-linguistic task had a significant difference between the proficiency groups. Contrary to what we expected, the results suggest that the resources recruited to store and judge the sequences of images were influenced by the proficiency level (differently from what was reported by Mota & Baltazar, 2015; Fontoura, 2018). On the other hand, Finardi (2009) found evidence that there is a correlation between WMC and L2 speech retention and acquisition of a complex syntactic structure, and the study of Azevedo (2012) indicates that WMC and L2 listening performance are related. Our results reveal that the native group was statistically different from the bilinguals in the non-linguistic task, and the higher proficiency group was also better than the lower proficiency group in the non-linguistic task. WMC may have aided participants to reproduce grammatical sentences, but proficiency was also important for participants to pay attention and store this linguistic cue. The added load caused by the non-linguistic performance did not seem to hinder the higher proficiency group's performance in detecting and correcting ungrammatical sentences. Nonetheless, we cannot forget to consider that the majority of our higher proficiency group consists of ESL teachers who have vast training in 3rd person singular morpheme. They deal with this kind of problem in their everyday routines, and they are used to correcting their students when failing to produce the 3rd person singular agreement.

CHAPTER 6

FINAL CONSIDERATIONS

In conclusion, our first hypothesis was confirmed. Grammatical sentences had a higher rate of expected answers than unexpected answers in all three groups (74% of expected answers by the lower proficiency group, 92% by the higher proficiency group, and 99% by the native group). The second hypothesis was not corroborated because the results within groups revealed that the rate of expected answers in local-distance and long-distance conditions were very similar. Lower proficiency participants did not exhibit a significant amount of correction of the ungrammatical sentences independently from the sentence condition, whereas higher proficiency participants corrected the four types of ungrammatical sentence conditions. The natives also had an equivalent performance in the target conditions. The problem was correcting ungrammatical sentences rather than reproducing grammatical sentences (ERLAM, 2006). Long-distance agreement sentences may be an extra challenge for lower proficiency learners (HOSHINO; DUSSIAS; KROLL, 2010; JENSEN et al., 2019), but these types of sentences were as problematic to correct as the local distance sentences were for our lower proficiency group.

The third hypothesis was partially confirmed because we found a marginal difference between the performance of higher proficiency and lower proficiency participants (z= -2.317, p= 0.0535) and a significant difference between the performance of natives and the lower proficiency participants (z=2.718, p= 0.0181) in the grammatical condition. WMC can be related to L2 learners' performance (FINARDI, 2009; AZEVEDO, 2012), and it could have had an effect on perceiving, storing, and reproducing the 3rd person singular morpheme (WEN; MOTA, MCNEILL, 2013).

We partially accept our fourth hypothesis because the lower proficiency group faced a challenge in correcting even local-distance ungrammatical sentences. The higher proficiency outperformed the lower proficiency group in correcting all target ungrammatical sentences with long-distance – animate (z= -2.317, p= 0.0066) and inanimate (z=-3.514, p= 0.0013) – and local-distance agreement – name (z=-3.321, p= 0.0026) and pronoun (z=-4.664, p= <0.0001). To our surprise, natives did not perform significantly differently than lower proficiency participants in the animate (z= 2.205, p= 0.0703) and the name (z= 1.644, p= 0.2272) conditions. Proficiency is important when L2 learners correct ungrammatical sentences (ERLAM, 2006); however, variability in different dialects may explain why natives failed to correct significantly more than lower proficiency participants in these conditions (LABOV et al., 1968). Proficiency was very important to define the L2 outcomes (HOSHINO; DUSSIAS; KROLL, 2010). Higher proficiency learners seemed to have developed more stages of the lexicon; therefore, they could retrieve and encode morphological information (JIANG, 2000; JACKENDOFF & AUDRING, 2016).

The fifth hypothesis was refuted. We found a significant difference between the performance of the higher proficiency and the lower proficiency groups ($z=8.690$, $p<0.0001$), between the native and the higher proficiency groups ($z= 3.135$, $p= 0.0049$), and the native and the lower proficiency groups ($z= 6.906$, $p<0.0001$) in image recall and judgment. WMC is very important to produce accurate performances, but it is also very important in rehearsal, noticing, attention regulation, and task-switching (WEN; MOTA; MCNEILL, 2013). Having to switch between sentence recall and image recall may have hindered the less proficient participants' attention to linguistic cues, such as the 3rd person singular morpheme presence and omission.

Souza (2014), Erlam (2006), Potter & Lombardi (1990, 1998), Lombardi & Potter (1992), Hamayan; Saegert & Larudee (1977), and Sachs (1967) conducted a sentence recall task in which participants had to orally reproduce or correct the sentences presented to them. We decided to adapt this task to be typed instead of orally produced because of the imposed difficulties of gathering data during the pandemic. This format did not allow us to test if participants could promptly correct ungrammatical sentences, but we could observe that lower proficiency participants had problems detecting written ungrammatical sentences in a time frame of 5,000 msec and correcting them in a time frame of 100,000 msec. Therefore, even if participants detected the violations, they were not able to correct them. The results suggest that proficiency was the most predominant factor to explain the problems in detecting and correcting ungrammatical sentences, but there could be other factors. Future researchers can also compare the results between a sentence recall task that is orally recalled and one that is recalled in a written format in order to check the different results that may arise from different task formats.

We had an intermediary memory task between sentence presentation and sentence recollection. This WM task could have added an extra load to the lower proficiency group. Memorizing the sentence in up to 5,000 msec, seeing the first image for 2,000 msec and the second for 2,000 msec, providing the image answer in up to 10,000 msec, and just then entering the sentence could have been too much for the lower proficiency group. Besides, there was also a 1,000 delay between each of these actions. We cannot forget that WMC is capacity-limited and time-limited (BADDELEY, 2003; COWAN, 2010); sentence violations could have passed unnoticed by lower proficiency participants after exhausting their WMC. Another possible interpretation is that some participants may have faced the task as a memory test. This could also explain why some natives reproduced typo and modal violations. Therefore, future research can assemble a task to test participants' ability to detect sentence violations with and without WM load.

Mota & Baltazar (2015) found no relation between WMC and morphological processing for L2 learners. They analyzed both regular and irregular past verb frequency, but we only worked with regular verbs. In the future, researchers revisiting the connection between WMC and local-distance and long-distance agreement with the 3rd person singular morpheme processing and production should have stimuli with both regular and irregular verbs to check whether these different types of verbs yield distinct processing and production difficulties from L2 learners.

L2 learners may present variability in their production. They may have acquired a form but may not be able to access it. Having a morphological representation does not entail that learners have the right morphological representation in their interlanguage (WHITE, 2003). Having explicit knowledge about the 3rd person singular morpheme does not mean that the learner will be able to use it (HAN, 2010, 2013). There is a big debate over the nature of explicit and implicit language knowledge and whether L2 learners can use explicit language knowledge in their production (cf. KRASHEN, [1981] 2002, [1982] 2009; DEKEYSER, 1998; R. ELLIS, 1994). Even in written tasks, where participants have time to monitor their production, it is not simple to place the 3rd person singular morpheme by the L2 learners with less expertise. Therefore, the schemas of lexical items can only be checked and employed if the knowledge is fully integrated into this shared network of representation and processing (JACKENDOFF & AUDRING, 2016). Rules and lists can co-exist for storing and generating lexical items (JACKENDOFF & AUDRING, 2016). They may start by being stored but are not erased from the lists once they can be generated (BOOJI, 2017). If they can be generated, they can also be stored. This type of redundancy makes computation more robust (LIBBEN, 2006). Learners have to figure out the similarities among items to store in this shared network of representation and processing in LTM. L2 learners may store verb forms individually, but there are links between them. This less automatic process may result in morphological mistakes in production (JIANG, 2000). Future studies can investigate the difference between explicit and implicit L2 language knowledge and how these types of knowledge are stored in LTM.

Another factor to be further investigated is low salience in the 3rd person singular morpheme. It can explain why this is such a troublesome form to be mastered. The 3rd person singular morpheme conveys some redundant information that can be retrieved by other elements in the sentence (GOLDSCHNEIDER & DEKEYSER, 2001; N. ELLIS, 2017; JENSEN et al., 2019). While L1 users use top-down mechanisms to process these forms, it is very hard for L2 learners to rely on bottom-up auditory processing to grasp these forms (N. ELLIS, 2017). Entrenchment is also a problematic issue in lexical development. While learners can balance it through resonance links between L1 and L2 (MACWHINNEY, 2017), it appears that L2 learners can only employ the morphological aspects of a lexical item if its development has been thoroughly reached (JIANG, 2000). Comprehending a lexical item is different from producing it (JIANG, 2000) because the learner needs vocabulary depth to produce it (SCHIMITT, 2014). The lexical item will trigger the speaker's message if it is stored in LTM (LEVELT, 1989). Comprehending and producing a low salient element such as the -s in 3rd person singular morpheme can be tough for lower proficiency participants. Future investigations can analyze the role of salience in L2 morpheme comprehension and production. It is also important to research the mechanisms used for morpheme comprehension and production in L1 and L2.

Many native speakers and some higher proficiency participants failed to correct ungrammatical sentences, and we cannot know for sure if they misunderstood the task. Erlam (2006) instructed participants that they should correct sentences' mistakes, and she also provided feedback on the sentences' outcome in the training session to avoid simple rote repetition from higher proficiency participants. Consequently, future researchers adopting this paradigm should have a version where they provide feedback in the training session on how sentences are expected to be corrected and another version without this feedback in the training session to check whether feedback on sentence correction results in different outcomes.

Even though we expected that participants placed the 3rd person singular morpheme where it was missing, some participants corrected the ungrammatical sentences to different grammatical versions of the sentence. This ranged from changes in time tense, lexical change, and so forth. Sachs (1967) argues that syntactic structures fade faster than the sentence meaning, but Lombardi & Potter (1992) and Potter & Lombardi (1998) claim that one structure can prime other structures.

When participants have such an open opportunity to rewrite the sentences, this type of unpredicted result can come out. Future researchers can adopt other paradigms such as the Maze Task (check Oliveira (2020), for instance) in order to force participants to choose between a grammatical and an ungrammatical option.

Despite the fact that we found no substantial difference between animate and inanimate conditions and between name and pronoun conditions in participants' production, Sachs (1967) defends that a sentence's meaning is longer retained than its form. This would also explain the tense and lexical changes found in participants' production. Thus, future sentence recall tasks could investigate the difference between semantic and structural devices to test which one would help to evoke higher correction rates.

We tried to assemble our stimuli as best as we could; thus, we controlled word frequency, syllables, and characters number. However, we failed to have the sentences checked by a native speaker beforehand because of time and personnel constraints. This is an issue that should be further looked into in the future.

It was very hard to find native participants for this study, particularly considering the Brazilian law of recruitment that prohibits researchers to pay participants to take part in their tasks (a common practice in many well-known and respected institutions around the world). Gathering data online during pandemic times was especially hard because of the nature of our task. Many participants found it tiring and gave up before completing it. This data could not be used, and we were left with a smaller number of participants than we would like to have, particularly with native speakers participants. Forthcoming research could be carried out in person and collect a more expressive number of participants from different proficiency levels. Considering the different levels of VLT, there could be a study with an adequate number of participants for each level to check at which level participants start detecting and producing the 3rd person singular morpheme.

References

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 Lingüística e Letras. Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2012.

BAAYEN, 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, 1992.

___________. Working memory and language: an overview. *Journal of Communication Disorders*, v. 36, p. 189-208, 2003.

BADDELEY, A.; GATHERCOLE, S.; PAPAGNO, C. The phonological loop as a language learning device. *Psychological Review,* n. 105, p. 158–73, 1998.

BADDELEY, A. D.; HITCH, G. Working memory. In: BOWER, G. H (Ed.), *The psychology of learning and motivation: Advances in research and theory*, New York: Academic Press, v. 8, p. 47–89, 1974.

BIALYSTOK, E.; FROHLICH, M. Second Language Learning and Teaching in Classroom Settings: The Learning Study, Year One. Toronto: Modern Language Centre, Ontario Institute for Studies in Education, p. 1-91, 1977.

BOCK, K & LEVELT, W. J. M. Language production: Grammatical encoding. In: GERNSBACHER, M. A. (Ed.) *Handbook of Psycholinguistics*. San Diego: Academic Press, 945-984, 1994.

BOOJI, G. *Construction morphology*. Oxford: Oxford University Press, 2010.

___________. Inheritance and motivation in Construction Morphology. In: GISBORNE & HIPPISLEY (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.

CARTON, A. S. Inferencing: A process in using and learning language. In: PIMSLEUR, P.; QUINN, P. (Eds.): *The Psychology of Second Language Learning*. Cambridge University Press, 1971.

CHOMSKY, N. *Aspects of the theory of syntax*. Cambridge: M.l.T. Press, 1965.

COWAN N. Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*. n. 104, p. 163-191, 1988.

___________. An embedded-processes model of working memory. In: MIYAKE, A.; SHAH, P. (Eds.): *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, NewYork: Cambridge University Press. p. 32-101, 1999.

___________. The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*. n. 19, p. 51-57, 2010.

DEKEYSER, R. Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In: DOUGHTY, C.; WILLIAMS, J. (Eds.), *Focus on Form in Classroom Second Language Acquisition*. Cambridge: Cambridge University Press, 1998.

DULAY, H.; BURT, M. Natural sequences in child second language acquisition. *Language Learning*, v. 24, p. 37-53, 1974.

DULAY, H; BURT, M. A new approach to discovering universal strategies of child second language acquisition. In: DATO, D. (Ed.) *Developmental Psycholinguistics: Theory and Applications.* Georgetown University Round Table on Languages and Linguistics. Washington: Georgetown University Press. p. 209-233, 1975.

ELLIS, N. C. Salience in Usage-based SLA. In: GASS, S.; SPINNER, P.; BEHNEY, J. (Eds.). *Saliency in Second Language Acquisition*. NY: Routledge, 2017.

ELLIS, R. *The Study of Second Language Acquisition*. Oxford: Oxford University Press, 1994.

___________. A Theory of Instructed Second Language Acquisition. In: ELLIS, N. (Ed.): *Implicit and explicit learning of Languages*. San Diego, CA, Academic Press, p. 79-114, 1994.

___________. Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching. In: ELLIS, R.; LOEWEN, S.; EDLER C.; ERLAM, R.; PHILP, J.; & REINDERS, H. (Eds.). (2009). *Implicit and explicit knowledge in second language learning, testing, and teaching*. Bristol, UK: Multilingual Matters.

ERLAM, R. Elicited Imitation as a Measure of L2 Implicit Knowledge: An Empirical Validation Study. *Applied Linguistics*, v. 27, n. 3, p. 464–491, 2006.

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.

FONTOURA, B. R. *A Study of English L2 Users' Processing Inflectional Morphemes and their Working Memory Capacity.* 2018. 78f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2018.

GODOY, M. C. Introdução aos modelos lineares mistos para os estudos da linguagem. *PsyArXiv*, 2019. https://doi.org/10.17605/OSF.IO/9T8UR

GOLDSCHNEIDER, J. M.; DEKEYSER, R. M. Explaining the "Natural Order of L2 Morpheme Acquisition" in English: a Meta Analysis of Multiple Determinants. *Language Learning*, v. 55. p. 27-62, 2001

HAMAYAN, E.; SAEGERT, J.; LARUDEE, P. Elicited imitation in second language learners. *Language and Speech*, v. 20, n.1, p. 86–97, 1977.

HAN, Z-H. Grammatical morpheme inadequacy as a function of linguistic relativity: A longitudinal study. In: HAN, Z.-H.; CADIERNO, T. (Eds). *Linguistic Relativity in SLA,* Clevedon, UK: Multilingual Matters, p. 154–182, 2010.

___________. Forty years later: Updating the Fossilization Hypothesis. *Language Teaching*, v. 46, n. 2, p. 133-171, 2013.

HOSHINO, N.; DUSSIAS, P. E.; KROLL, J. K. Processing subject-verb agreement in a second language depends on proficiency. *Bilingualism: Language and Cognition*, v. 13, p. 87-98, 2010.

HUDSON, R. *An Introduction to Word Grammar*. New York: Cambridge University Press, 2010.

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

JENSEN, I.; SLABAKOVA, R.; WESTERGAARD, M.; LUNDQUIST, B. The Bottleneck Hypothesis in L2 acquisition: L1 Norwegian learners' knowledge of syntax and morphology in L2 English. *Second Language Research*, v. 36, n. 1, p. 3-29, 2019.

JIANG, N. Lexical representation and development in a second language. *Applied linguistics*, v. 21, n. 1, p. 47-77, 2000.

___________. 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.

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.

KRASHEN, S. Some issues relating to the Monitor Model. In: BROWN, H. D.; YORIO, C.; CRYMES, R. (Eds.). *Teaching and Learning English as Second Language: On TESOL '77*. Washington: TESOL, p. 144-158, 1977.

KRASHEN, S. *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon, [1981] 2002.

___________. *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon, [1982] 2009.

KRASHEN, S., HOUCK, N., GIUNCHI, P, BODE, S., BIRNBAUM, R.; STREI, J. Difficulty order for grammatical morphemes for adult second language performers using free speech. *TESOL Quarterly,* v. 11, p. 338-341, 1977.

LABOV, W.; COHEN, P.; ROBBINS, C; LEWIS, J. A *Study of the Non-Standard English of Negro and Puerto-Rican Speakers in New York City. Final Report*, Cooperative Research Project 3228, Philadelphia: US Regional Survey, 1968.

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

LAUFER, B.; NATION, P. Passive vocabulary size and speed of meaning recognition: Are they related? *EUROSLA Yearbook*, v. 1, p. 7-28, 2001.

LEVELT, W. J. M. *Speaking: From Intention to Articulation*. Cambridge, MA: Bradford, 1989.

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

LOMBARDI, L.; POTTER, M. C. The regeneration of syntax in short term memory. *Journal of Memory and Language*, v. 31, n. 6, p. 713-733, 1992.

MACWHINNEY, B. Entrenchment in second-language learning. In: SCHMID, H.-J. (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge,* p. 343–366, 2017.

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, 2008.

MILLER G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, v. 63, p. 81–97, 1956.

MOTA, M. B.; BALTAZAR, L. M. O processamento da morfologia flexional na L2: efeitos de frequência, proficiência, controle inibitório e memória de trabalho. *Prolíngua*, v. 10, n.1, p. 134-149, 2015.

MOTA, M. B.; KRAMER, R. Effects of bilingualism on inhibitory control and working memory: A study with early and late bilinguals. *Gragoatá*, v. 20, n.38, p. 309-331, 2015.

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

OLIVEIRA, C. S. F. Métodos on-line em Psicolinguística: A Tarefa Labirinto (Maze Task). *CADERNOS DE TRADUÇÃO*, v. 40, p. 217-248, 2020.

OLIVEIRA, C. S. F; FONTOURA, B. R.; SOUZA, R. A. The challenges imposed by L2 inflectional morphology: evidence from speeded acceptability judgment tasks with Brazilian Portuguese-English bilinguals. *Veredas: Revista de Estudos Linguísticos,* v. 24. n. 1, p. 317-339, 2020.

POTTER, M.; LOMBARDI, L. Regeneration in the short term recall of sentences. *Journal of Memory and Language,* v. 29, p. 633–54, 1990.

___________. Syntactic priming in the immediate recall of sentences. *Journal of Memory and Language*, v. 38, n. 3, p. 265-282, 1998.

SACHS, J. S. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, v. 2, p. 437-442, 1967.

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.

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

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. What is easy and what is hard to acquire in a second language? In: BOWLES, M. et al. (Eds.). *Proceedings of the 10th generative approaches to second language acquisition conference*. Somerville, MA: Cascadilla Proceedings Project, p. 280-294, 2013.

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

SOUZA, R. A. Memorização de sentenças com construção específica de uma das línguas de bilíngues. *Letrônica*, v. 7, 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.; 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.

STOET, G. PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, v. 42, n.4, p. 1096-1104, 2010.

STOET, G. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, v. 44, n.1, p. 24-31, 2017.

WHITE, L. *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press, 2003.

ZOBL, H.; LICERAS, J. Functional categories and acquisition orders. *Language Learning,* v. 44, p. 159–180, 1994.

APPENDICES

Appendix 1

Trial

letter_A.jpg	letter_B.jpg
armchair_1.jpg	armchair_2.jpg
letter_a_lower.jpg	letter_a_lower.jpg
balloon.jpg	balloon.jpg
letter_C_1.jpg	letter_C_2.jpg
cat_1.jpg	cat_2.jpg
letter_b_lower.jpg	letter_b_lower.jpg
cookie.jpg	cookie.jpg
letter_E.jpg	letter_F.jpg
city_1jpg	city_2.jpg
letter_e_lower.jpg	letter_e_lower.jpg
stove.jpg	stove.jpg
letter_D_1.jpg	letter_D_2.jpg
door_1.jpg	door_2.jpg
letter_f_lower.jpg	letter_f_lower.jpg
statue.jpg	statue.jpg
letter_G.jpg	letter_H.jpg
forest_1	forest_2
letter_g_lower.jpg	letter_g_lower.jpg
zebra.jpg	zebra.jpg
letter_I_1.jpg	letter_I_2.jpg
house_1.jpg	house_2.jpg
letter_h_lower.jpg	letter_h_lower.jpg
nailpolish.jpg	nailpolish.jpg
letter_J.jpg	letter_K.jpg
tree_1.jpg	tree_2.jpg
letter_j_lower.jpg	letter_j_lower.jpg
elephant.jpg	elephant.jpg

letter_L_1.jpg	letter_L_2.jpg
book_1.jpg	book_2.jpg
letter_k_lower.jpg	letter_k_lower.jpg
flowers.jpg	flowers.jpg
letter_N.jpg	letter_O.jpg
snow_1	snow_2
letter_n_lower.jpg	letter_n_lower.jpg
fish.jpg	fish.jpg
letter_M_1.jpg	letter_M_2.jpg
pool_1.jpg	pool_2.jpg
letter_o_lower.jpg	letter_o_lower.jpg
classroom.jpg	classroom.jpg
letter_P.jpg	letter_Q.jpg
farm_1	farm_2

Appendix 2[62]

Welcome Screen

Dear participant, this task is part of the Ph.D. research in Linguistics (PosLin / UFMG) by Bruna Rodrigues Fontoura.

If you are between 18 and 35 years old and an English native speaker, you can perform this task.

Moreover, participants who have any kind of color blindness cannot carry out this task.

All your personal data collected will be kept secret.

This task demands that English sentences and images be memorized.

We estimate that this task takes between 20 to 30 minutes to be completed

Please, do this task with maximum attention and concentration.

Thank you very much for your interest and willingness to contribute.

Screen 1

Dear participant, you will take part in a sentence recall task with images.

You will see a sentence that must be memorized, and then you will see two images and say whether they are the same or not. Afterwards, you will type the sentence.

Press the spacebar to continue

Screen 2

You must do the best you can, as fast as you can.

Therefore, you will:
- see the sentence and memorize it;
- say whether the images are the same or not;
- type the sentence.

---

[62] Instructions were provided in the participants' native language.

Press the spacebar to continue

Screen 3

It is important to remember that the images must have the same shape and color to match.

To judge whether the images match, you should press:

y = yes, they are the same;
n = no, they are not the same;

And then press the "enter" key.

Press the spacebar to continue

Screen 4

After judging the images, you will see where to type the sentence.

If you identify a mistake in the sentence, you must rewrite it in the correct form.

If you want to correct any character, you must press the key corresponding to "backspace" on your keyboard.

It is not necessary to put a period at the end of sentences.

Press the spacebar to continue

Screen 5

Even if you are not sure, you must type the sentence and, right after, press the "enter" key.

Do not leave blank sentences.

By typing the sentence and pressing "enter" you will proceed to the next screen.

Press the spacebar to continue

Screen 6

Let's practice the task with some examples?

In the training session, you will receive feedback on your performance in judging the images.

You will see when your performance is correct, incorrect, or has to be faster.

Press the spacebar to continue

Screen 7

Recalling the steps:

- Memorize the sentence displayed

- Judge the images:
y = yes, they are the same
n = no, they are not the same
enter

- Type the sentence in the correct form
enter

Press to start the training session

End of training

This is the end of the training session.

When you're ready, press the spacebar to start the task.

End of Experiment

This is the end of the task. Thank you very much for your participation!

Press the spacebar to end

Appendix 3

(1) **Question for natives before the experiment**

1. What's your full?

2. What's your email?

3. How old are you?

4. What's your occupation?

5. Where are you from? (City and Country)

6. Where do you live currently? (City and Country)

7. Do you speak another language? If you do, which one? How old were you when you started learning it?

8. What's your current educational level?

- Complete high school.

- Undergraduate course in progress.

- Complete undergraduate course.

- Graduate course in progress.

- Complete graduate course.

9. Do you have any type of color blindness?

- yes.

- no.

(2) **Questions for L2 learners before the experiment**

1. What's your full?

2. What's your email?

3. How old are you?

4. What’s your occupation?

5. Where are you from? (City and Country)

6. Where do you live currently? (City and Country)

7. How much time did you study English?

8. How old were you when you started learning English?

9. Have you ever lived in an English speaking country? Where? For how long?

10. What's your current educational level?

- Complete high school.

- Undergraduate course in progress.

- Complete undergraduate course.

- Graduate course in progress.

- Complete graduate course.

11. Do you have any type of color blindness?

- yes.

- no.

(3) **Common questions after the experiment**

A. How well do you type?

- I don’t type well.

- I type somewhat well.

- I type well.

- I type very well.

B. How fast do you type?

- I don't type fast

- I type somewhat fast.

- I type fast.

- I type very fast.

Appendix 4

FREE AND INFORMED CONSENT TERM[63]

CAAE Project: 46886921.8.0000.8507, approved by the CEP/CONEP System, on August 12, 2021.

We invite you to participate in an experiment that will comprise research coordinated by Prof. Dr. Cândido Samuel Fonseca de Oliveira (MG 10,855,750), English language teacher at the Federal Center for Technological Education of Minas Gerais (CEFET-MG). This invitation is due to the fact that you are a Brazilian Portuguese speaker and an English language learner or an English native speaker, which is very useful for our research. This research belongs to a study on the applicability of a language learning technique. The survey consists of a few steps: a brief questionnaire about your personal information (e.g. age, native language, and education) and the diagnosis of psychiatric illness or disability, a proficiency test, a language profile questionnaire, and, finally, the language tests in Portuguese or English. In the language tasks, you will perform tasks such as (i) rating from 1 to 5 how natural the grammatical structures of each sentence sound (acceptability judgment), (ii) completing sentences (cloze), (iii) reading sentences word by word (self-paced reading), (iv) assembling sentences word by word according to the options offered (maze task) and/or (v) memorizing sentences and images and typing the sentence that was memorized (sentence recall task). The invitation you received explains to which of these steps you are being invited. The questionnaires help us to understand your profile as a speaker of one or more languages and the tasks provide us with important information about the processing and learning of different linguistic structures. You may discontinue your participation in the study at any time, without penalty or prejudice. You will be given instructions for all tasks before performing them. All data collected will be converted into numerical and quantitative information that will be accessed exclusively by the researchers of this study. Thus, all participants will be guaranteed anonymity, that is, under no circumstances will your identity as a participant be revealed. Each task lasts approximately 20 minutes with at least one break for you to rest. Tasks require concentration, you can take your time, and, as risks, you may feel slightly tired. In addition, there is a risk of data leakage, as the collection takes place virtually, but there is a preparation for this not to occur, such as choosing to use password-protected platforms to carry out the collection. However, your participation does not involve any form of risk to your physical and mental health, nor may it compromise your reputation as a citizen, as a student, or as a professional. Your participation in this study will not be of immediate benefit to you individually. The results of this study may be made public in the form of oral or poster presentations at scientific conferences; scientific articles published in specialized journals; book chapters; monographs, dissertations, or theses guided by the responsible researcher. However, for the reasons listed above the identities of the participants will remain hidden in any and all forms of publication and dissemination of the study.

As a research participant and in accordance with Brazilian legislation, you have several rights, in addition to anonymity, confidentiality, secrecy, and privacy, even after the research has ended or been interrupted. Thus, you are guaranteed:

- Compliance with the practices determined by the applicable legislation, including Resolutions 466 (and, in particular, its item IV.3) and 510 of the National Health Council, which regulate research ethics and this Term;

[63] Participants had access to the free and informed consent term in their native language.

- Full freedom to decide on their participation without prejudice or reprisal of any nature;
- The full freedom to withdraw your consent, at any stage of the research, without prejudice or reprisal of any nature. In this case, data collected from your participation up to the time of withdrawal of consent will be discarded unless you explicitly authorize otherwise;
- Monitoring and assistance, even if subsequent to the termination or interruption of the research, free of charge, in full and immediately, for the necessary time, whenever required and related to their participation in the research, upon request to the responsible researcher;
- Access to search results;
- Reimbursement of any expenses related to participation in the research (for example, travel costs to the agreed place for the interview), including any accompanying person, upon request to the responsible researcher;
- Compensation for possible damages resulting from the research;
- Access to this Term. This document is initialed and signed by you and a researcher from the research team, in two copies, one copy of which will remain your property. If you lose your copy, you can still request a copy of the document from the responsible researcher.

If you have any questions, feel free to contact the responsible researcher. You can also consult the CEP (Ethics and Research Committee of the Federal Center for Technological Education of Minas Gerais) to ask questions about ethical aspects of research.

Responsible Researcher:
Prof. Dr. Cândido Samuel Fonseca de Oliveira
General Training Department of the Contagem Unit
Federal Center for Technological Education of Minas Gerais
coliveira@cefetmg.br - (31) 3368-4300

Research Ethics Committee(CEP):
Campus VI
Av. Amazonas, 5855, Gameleira.
Belo Horizonte – MG. Zip code: 30.510-000
Phone: +55 (31) 3379-3004

____
Cândido Samuel Fonseca de Oliveira

If you agree to take part in the research, please complete and sign the consent form below.
We thank you for your participation!

CONSENT FORM:

I, ____, agree to participate in the task that makes up the experimental study on English learning supervised by Prof. Dr. Cândido Samuel Fonseca de Oliveira from the Federal Center for Technological Education of Minas Gerais. I am aware that my participation in the study does not imply harm or risk to my health and well-being, which does not imply risks to my reputation and image, and that it does not imply immediate benefits for me, individually. I am also aware that I can discontinue my participation in the experiment at any time, free of charge, and that my identity as a participant will not be exposed in any form of disclosure of the results of the study.

Date:
Participant: ____

Appendix 5

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 6

In this section, we decided to include the script that we used in our task. We tried to comment (after the # symbol) as much as possible so that it could be some sort of guidelines for other people in need. Everyone that uploads the images and the sentences files can actually replicate this same task. We hope that this can be helpful for future research, especially for those dealing with images memorization and typed answers.

Psytoolkit Script

```
include sentence.txt #sentences file that is going to be displayed

options
        bitmapdir images #images file that is going to be displayed
        set &waitTime 2000 #time of image display in milliseconds
        set &intertime 1000 #time between displays in milliseconds
        set &waitSentenceTime 5000 #time of sentence display in milliseconds
        set &numTraining 4 #number of sequences displayed in training (sentence + image 1+
image 2)
        set &numTrial 42 #number of sequences displayed in trial (sentence + image 1+ image 2)

        set &total 46 #total number of sequences displayed (sentence + image 1+ image 2 in
training + trial)
        fullscreen #force fullscreen
        resolution 1100 800 #resolution format
        background color white #background color

fonts
        my_arial arial 20 #font and size of letters displayed whenever participants type something

bitmaps
        include bitmaplist.txt #bitmap file that includes the sequence of images that is going to be
displayed (including instructions that are saved as images)

part showSentence #sentence display
        set $sentence &&sentenceList[&firstImageIndex] #sentence displayed taken from a
position on the list. It shuffles and picks one
        text color black #color of sentence displayed
        show text $sentence #shows the picked sentence
```

part delaySentenceTime #check part delaySentenceTime below
##index(identifier = variable). Each sentence has a number that identifies it, namely, index

part showFirstImage #1st image display
set &firstImageIndex &&randomIndexList[&trialCount] #in order to have a random list for the images, we had to create a list. This list has the images indexes (images number) as its content. We created the list and shuffled it. However, every time it is randomized[64], it needs to look for the second corresponding image (or value) for the interaction, otherwise, there would be chaos.
show bitmap &firstImageIndex #shows the 1st image from the pair
part delayTime #check part delayTime below

part showSecondImage #2nd image display
set $secondImage expression &firstImageIndex + &total #the second image is displayed after the first image index; therefore, it shows the corresponding pair. + &total: the total of pairs. We have 46 pairs (training +trial). If we add up to the total, we have the corresponding pair
show bitmap $secondImage #shows 2nd image on the screen
part delayTime
part getImageComparisonUserAnswer #image answer entered by participant
part checkImageAnswer #checks if the participant entered the right answer for image display
part delayIntertime #check part delayIntertime below

part getImageComparisonUserAnswer #image answer entered by participant
text color black #color of answer typed by participant
readkeys option font my_arial #readkeys: waits for user typed answer
readkeys option size 100 # size: space reserved for participants to type the answers, 100 means the amount of characters space available to type, you can put as many as you want to
readkeys option space 17 #space between the letters (less than 17 gets cramped)
show bitmap image_box #background image used on the image answer
readkeys option show 0 0 #position on the screen (screen center)
if &&checklist[&firstImageIndex] == 1 #answer validation to whether participants entered the right answer
readkeys "y" 10000 0 100 #time waiting for user "y"[65] answer in up to 10000 milliseconds
else
readkeys "n" 10000 0 100 #time waiting for user "n" answer in up to 10000 milliseconds

64 Randomize the list: the tool enables you to do it that way, and not in a better way.
65 "s" for the Portuguese version.

```
        fi
        set &imageRT expression RT #saves image answer RT
        clear screen

part getUserInputSentence #sentence typed by participant
        text color black #color of answer typed by participant
        readkeys option font my_arial #readkeys: waits for user typed answer
        readkeys option size 100 # size: space reserved for participants to type the answers, 100
means the amount of characters space available to type
        readkeys option space 17 #space between the letters (less than 17 gets cramped )
        show bitmap empty_box #background image used on the sentence answer
        readkeys option show -450 0  #position on the screen
        readkeys &&sentenceList[&firstImageIndex] 100000 0 100 #time waiting for user typed
answer in 100000 milliseconds
        set &sentenceRT expression RT #saves sentence RT
        clear screen

part delayTime
        delay &waitTime #2000: time of image display in milliseconds
        part delayIntertime

part delaySentenceTime
        delay &waitSentenceTime #5000: time of sentence display in milliseconds
        part delayIntertime

part delayIntertime
        clear screen
        delay &intertime #1000: time between displays in milliseconds

part showGetReady
        show bitmap ready #message displayed before each new sequence for participant to get
ready
        part delayTime

part checkSentenceAnswer #check sentence answer
        if STATUS == TIMEOUT #timeout in sentence typed by participant
                set &is_correct_sentence_answer "no" #if there is no typed answer in 100000
milliseconds, participants see a message to go faster
                part showTimeoutBitmap #message to go faster
        fi
        if STATUS == CORRECT #the sentence matches the one presented to participant
```

```
        set &is_correct_sentence_answer "yes" #typed answer matches (saves the answer
entered by participants)
    fi
    if STATUS == WRONG #the sentence does not match the one presented to participant
        set &is_correct_sentence_answer "no" #typed answer does not match(saves the
answer entered by participants)
    fi

part checkImageAnswer #check image answer
    set &miss "no" #This is a flag when participants provide no answer. First of all, we
consider that the person will enter an answer and then we consider the other scenarios
    if STATUS == TIMEOUT #timeout if no answer is provided in 10000  milliseconds
        set &miss "yes" #when no answer is provided, we have a missed answer
        set &is_correct_image_answer "no" #when there is a timeout, it records as an
incorrect answer, putting no to the row is_correct_image_answer
        set &image_answer "missed" #when there is a timeout, it records as a missed
answer
        if &&checklist[&firstImageIndex] == 1 #checks the list for the expected answer
            set &expected_image_answer "match" #expected match answer
        else
            set &expected_image_answer "mismatch" #expected mismatch answer
        fi
        part showTimeoutBitmap #shows a message to go faster
    fi
    if STATUS == CORRECT #correct image answer
        set &is_correct_image_answer "yes" #records it as a correct answer
        if &&checklist[&firstImageIndex] == 1 #checks the list for the image answer
            set &expected_image_answer "match"  #expected match answer
            set &image_answer "match"  #match answer provided by user
        else
            set &expected_image_answer "mismatch" #expected mismatch answer
            set &image_answer "mismatch" #mismatch answer provided by user
        fi
        part showCheckBitmap #shows a message of correct answer
    fi
    if STATUS == WRONG #wrong image answer
        set &is_correct_image_answer "no" #records it as a wrong answer
        if &&checklist[&firstImageIndex] == 1 # checks the list for the image answer
            set &expected_image_answer "match" #expected match answer
            set &image_answer "unmatch" #unmatch answer provided by user
        else
```

```
                set &expected_image_answer "mismatch" #expected mismatch answer
                set &image_answer "match" #match answer provided by user
            fi
            part showWrongBitmap #shows a message of incorrect answer
    fi

part showCheckBitmap #image displayed in correct image answer
    if BLOCKNAME == "training" #message displayed in training session
            show bitmap check #shows a message of correct answer
            part delayTime
    fi

part showWrongBitmap #image displayed in wrong image answer
    if BLOCKNAME == "training" #message displayed in training session
            show bitmap wrong #shows a message of incorrect answer
            part delayTime
    fi

part showTimeoutBitmap #image displayed in no image answer
    if BLOCKNAME == "training"  #message displayed in training session
            show bitmap timeout  #shows a message to go faster
            part delayTime
    fi

task initTable #task commands from file to code
    table table1 #loads the file table
    set &index @1 #retrieves the first-row value and puts it on the index position (each item, variable, has a value to identify itself called index)
    set &&sentenceList[&index] @2 #retrieves the second-row value and puts it on the list called sentenceList on the index position
    set &&checklist[&index] @3 #retrieves the third-row value and puts it on the list called checklist on the index position
##@1 @3 @2: rows for each separated value, @1 gets the value from row 1, @2 from 2, and @3 from 3, each interaction gets one line from the corresponding row

task oneback #order of the task
    set &trialCount increase  #It counts the interactions and starts each new interaction (sequence). Consequently, it picks one image and finds the second image from the pair. It starts at 0, adds 1, and goes on to wherever you set (in our case, 46 pairs). It increases the value in 0 by 1. It is the reference index to go on in the list. 0: standard, the starting element of the sequence. +1: it adds up 1 to this variable. It goes over every single item in the list in a random way.
```

set &firstImageIndex &&randomIndexList[&trialCount] #gets the value (firstImageIndex) that was put on the list on the last step and finds the second image (trialCount). The trialCount is sequential and used as an index to the random list (randomIndexList = random image list). Every list has an imaginary index that represents the element position on the list. It is going over the random elements on the list one by one. We set the list considering the image pairs (46 pairs)

part showGetReady #shows get ready image
part showSentence #shows sentence
part showFirstImage #shows 1st image
part showSecondImage #shows 2nd image
part getUserInputSentence #waits for typed answer
part checkSentenceAnswer #checks answer
save &trialCount ";" BLOCKNAME ";" &miss ";" &is_correct_image_answer ";" &expected_image_answer ";" &image_answer ";" &imageRT ";" &is_correct_sentence_answer ";" &&sentenceList[&firstImageIndex] ";" WORD ";" &sentenceRT ";" #names of things saved to be analyzed in a table

block init
tasklist #describes which task is being called
initTable 46 fixed #total number of sequences (sentence + image 1+ image 2 in training + trial)
end

block training #training block (the block is the first place the code goes over)
set &&randomIndexList clear #clears the random list. It clears the content from the variables
set &&randomIndexList range 1 &numTraining #sets a list that has the numbers from 1 and numTraining (number of sequences in Training = 4)
set &&randomIndexList shuffle #shuffles the list order
set &trialCount 0 #starts the trialCount here
message instr_1   #instruction messages that are listed in bitmap list and displayed to participant
message instr_2
message instr_3
message instr_4
message instr_5
message instr_6
message instr_7
tasklist #describes which task is being called
oneback 4 fixed #total number of training sequences (sentence + image 1+ image2)

end

block trial #trial block

set &startindex expression &numTraining + 1 #The trial starts here. Since we have the same list for the training and the trial, we need to establish that the first index for the trial starts after the numTraining (number of sequences in training, namely, 4). The trial goes from 5 to 46

set &endindex expression &numTraining + &numTrial #The trial ends here after 4 numTraining + 42 &numTrial

set &&randomIndexList clear #clears random list

set &&randomIndexList range &startindex &endindex #ranges from the starting to the ending point (from 5 to 46 in the trial)

set &&randomIndexList shuffle #shuffles the list order

set &trialCount 0 #the trialCount starts here

message instr_end_train #message displayed of training end

tasklist

oneback 42 fixed #total number of trial sequences (sentence + image 1+ image 2)

end

message instr_end_exp #message end of experiment

Appendix 7

Participants SB and KT provided some feedback that helped us understand a little more about their answers in the modal condition. SB even let us know that there were some broken English sentences in the task[66]:

Table SB: Modal Answers

Participant SB saw:	**Participant SB entered:**
Sam will to tell jokes	Sam will to tell jokes (not correct english.)
Liam will to kill the bird	Liam will kill the bird
Eve will to ride a bike	Eve will ride a bike
Claire will to make a mess	Claire made a mess
Fred will to pay the bill	Fred will pay the bill
Joe will to play football	Joe will play football

Table KT: Modal Answers

Participant KT saw:	**Participant KT entered:**
Sam will to tell jokes	Sam will to" tell jokes (Idk how to fix this one)"
Liam will to kill the bird	Liam will to" kill the birds (also don't know how to fix)"
Eve will to ride a bike	Eve will ride a bike (I think I get how to fix it now)
Fred will to pay the bill	Fred will pay the bill
Claire will to make a mess	Claire will make a mess
Joe will to play football	Joe will play football

66 Even though the answers are compiled here for the same condition, the script for the task generated a new random list of sequences for each participant.

Appendix 8

Script for the R analysis based on Godoy (2019):

```
#Load packages
library(dplyr)
library(ggplot2)
library(lme4)
library(lmerTest)
library(nlme)
library(emmeans)

##############################SENTENCES#####################################

performance_r_groups <- read_csv("Desktop/analysis/final/final-
analysis/performance_r_groups.csv")
View(performance_r_groups)

performance_r_groups <- read_csv("performance_r_groups.csv",
                        col_types = cols(VLT = col_character()))
View(performance_r_groups)

#Inspect the first lines
head(performance_r_groups)

#Check the variables
str(performance_r_groups)

#Let's check the values at "grammatical"
unique(performance_r_groups$grammatical)
#Let's check the values at "typos"
unique(performance_r_groups$typo)
#Let's check the values at "modal"
unique(performance_r_groups$modal)
#Let's check the values at "ungrammatical_pronoun"
unique(performance_r_groups$ungram.pronoun)
#Let's check the values at "ungrammatical_name"
unique(performance_r_groups$ungram.name)
#Let's check the values at "ungrammatical_animate"
unique(performance_r_groups$ungram.animate)
#Let's check the values at "ungrammatical_inanimate"
```

```
unique(performance_r_groups$ungram.inanimate)
#Let's check the values at "VLT"
unique(performance_r_groups$VLT)
#Let's check the values at "item"
unique(performance_r_groups$item)
#Let's check the values at "participant"
unique(performance_r_groups$participant)

#Expected answer  = correct answer
#Unexpected answer = incorrect
#NA = incomplete and blank

#filter NA answer
performance.sent.edit = performance_r_groups%>%
  select (grammatical, modal, typo, ungram.pronoun, ungram.name, ungram.animate,
ungram.inanimate, VLT)%>%
  filter(grammatical != "NA")%>%
  filter(ungram.name != "NA")%>%
  filter(ungram.animate != "NA")%>%
  filter(ungram.inanimate != "NA")%>%
  droplevels()

#table with the relative frequency of expected/unexpected answers in VLT in the grammatical
condition
performance.sent.edit%>%
  group_by(VLT, grammatical)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#grammatical graph
performance.sent.edit%>%
  group_by(VLT, grammatical)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = VLT, y = frequency, fill = grammatical))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("paleturquoise3", "salmon"))

#table with the relative frequency of expected/unexpected answers in VLT in the typo condition
```

```
performance.sent.edit%>%
  group_by(VLT, typo)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#typo graph
performance.sent.edit%>%
  group_by(VLT, typo)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = VLT, y = frequency, fill = typo))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("aquamarine3", "salmon"))

#table with the relative frequency of expected/unexpected answers in VLT in the modal condition
performance.sent.edit%>%
  group_by(VLT, modal)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#modal graph
performance.sent.edit%>%
  group_by(VLT, modal)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = VLT, y = frequency, fill = modal))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("aquamarine3", "salmon"))

###################################RATERS###################################

performance_groups_raters <- read_csv("performance.groups.raters.csv")
View(performance_groups_raters)

#Checking variables
#condition
unique(performance_groups_raters$condition)
```

```
#proficiency
unique(performance_groups_raters$proficiency)
#item
unique(performance_groups_raters$item)
#participant
unique(performance_groups_raters$participant)
#answer
unique(performance_groups_raters$answer)

#LOWER

#filter = lower group and the NA answers
perfomance.lower = performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(proficiency == "lower proficiency")%>%
  filter(answer != "NA")%>%
  droplevels()

#table with the relative frequency of expected/unexpected answers in the lower proficiency group
for each condition
perfomance.lower%>%
  group_by(condition, answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#lower graph
perfomance.lower%>%
 group_by(condition, answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = condition, y = frequency, fill = answer))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("paleturquoise3", "plum3"))

#model
model.lower = glm(as.factor(answer) ~ as.factor(condition), data = perfomance.lower, family =
binomial)
summary(model.lower)
```

```
#does not converge
model.lower1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                        (1+condition|item) + (1+condition|participant),
                      data = perfomance.lower, family = binomial(link = "logit"),
                      control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.lower1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                       (1+condition|item) + (1|participant),
                     data = perfomance.lower, family = binomial(link = "logit"),
                     control = glmerControl(optimizer = "bobyqa"))
#model 3 = works
model.lower1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                       (1|item) + (1|participant),
                     data = perfomance.lower, family = binomial(link = "logit"),
                     control = glmerControl(optimizer = "bobyqa"))

summary(model.lower1)

#model null
model.lower.null <- glmer(as.factor(answer)  ~ 1 +
                       (1|item) + (1|participant),
                     data = perfomance.lower, family = binomial(link = "logit"),
                     control = glmerControl(optimizer = "bobyqa"))

summary(model.lower.null)

anova(model.lower1, model.lower.null)

###POST-HOC for lower

post.hoc = emmeans(model.lower1, ~ condition)

pairs(post.hoc, adjust="tukey")

#HIGHER

#filter = higher proficiency group and NA answers
perfomance.higher = performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(proficiency == "higher proficiency")%>%
```

```
  filter(answer != "NA")%>%
  droplevels()

#table with the relative frequency of expected/unexpected answers in the higher proficiency
group for each condition
perfomance.higher%>%
  group_by(condition, answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#higher graph
perfomance.higher%>%
  group_by(condition, answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = condition, y = frequency, fill = answer))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("paleturquoise3", "plum3"))

#model
model.higher = glm(as.factor(answer) ~ as.factor(condition), data = perfomance.higher, family =
binomial)
summary(model.higher)

#does not converge
model.higher1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                        (1+condition|item) + (1+condition|participant),
                      data = perfomance.higher, family = binomial(link = "logit"),
                      control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.higher1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                        (1+condition|item) + (1|participant),
                      data = perfomance.higher, family = binomial(link = "logit"),
                      control = glmerControl(optimizer = "bobyqa"))
#model 3 = works
model.higher1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                        (1|item) + (1|participant),
                      data = perfomance.higher, family = binomial(link = "logit"),
```

```
                control = glmerControl(optimizer = "bobyqa"))

summary(model.higher1)

#model null
model.higher.null <- glmer(as.factor(answer)  ~ 1 +
                     (1|item) + (1|participant),
                   data = perfomance.higher, family = binomial(link = "logit"),
                   control = glmerControl(optimizer = "bobyqa"))

summary(model.lower.null)

anova(model.higher1, model.higher.null)

#POST-HOC for higher
post.hoc = emmeans(model.higher1, ~ condition)

pairs(post.hoc, adjust="tukey")

#NATIVE

#filter = native group and NA answers
perfomance.native = performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(proficiency == "native")%>%
  filter(answer != "NA")%>%
  droplevels()

#table with the relative frequency of expected/unexpected answers in the native group for each
condition
perfomance.native%>%
  group_by(condition, answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#native graph
perfomance.native%>%
  group_by(condition, answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = condition, y = frequency, fill = answer))+
```

```
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("paleturquoise3", "plum3"))

#model
model.native = glm(as.factor(answer) ~ as.factor(condition), data = perfomance.native, family =
binomial)
summary(model.native)

#does not converge
model.native1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                        (1+condition|item) + (1+condition|participant),
                      data = perfomance.native, family = binomial(link = "logit"),
                      control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.native1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                        (1+condition|item) + (1|participant),
                      data = perfomance.native, family = binomial(link = "logit"),
                      control = glmerControl(optimizer = "bobyqa"))
#model 3 = works
model.native1 <- glmer(as.factor(answer)  ~ as.factor(condition) +
                        (1|item) + (1|participant),
                      data = perfomance.native, family = binomial(link = "logit"),
                      control = glmerControl(optimizer = "bobyqa"))

summary(model.native1)

#model null

model.native.null <- glmer(as.factor(answer)  ~ 1 +
                            (1|item) + (1|participant),
                          data = perfomance.native, family = binomial(link = "logit"),
                          control = glmerControl(optimizer = "bobyqa"))

summary(model.lower.null)

anova(model.native1, model.native.null)

#POST-HOC for native
```

```
post.hoc = emmeans(model.native1, ~ condition)

pairs(post.hoc, adjust="tukey")

######COMPARISON BETWEEN GROUPS

#grammatical

#filter grammatical condition and NA answers
performance.grammatical = performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(condition == "grammatical")%>%
  filter(answer != "NA")%>%
  droplevels()

#table with relative frequency for the grammatical condition
performance.grammatical%>%
  group_by(proficiency, answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#grammatical graph
performance.grammatical%>%
  group_by(proficiency, answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = proficiency, y = frequency, fill = answer))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("cadetblue2", "lightpink3"))


model.grammatical = glm(as.factor(proficiency) ~ as.factor(answer), data =
performance.grammatical, family = binomial)
summary(model.grammatical)

#does not converge
model.grammatical1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                        (1+proficiency|item) + (1+proficiency|participant),
                       data = performance.grammatical, family = binomial(link = "logit"),
```

```
                    control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.grammatical1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                      (1+proficiency|item) + (1|participant),
                    data = performance.grammatical, family = binomial(link = "logit"),
                    control = glmerControl(optimizer = "bobyqa"))
#model 3 = works
model.grammatical1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                      (1|item) + (1|participant),
                    data = performance.grammatical, family = binomial(link = "logit"),
                    control = glmerControl(optimizer = "bobyqa"))

summary(model.grammatical1)

#model null
model.grammatical.null <- glmer(as.factor(answer)  ~ 1 +
                          (1|item) + (1|participant),
                         data = performance.grammatical, family = binomial(link = "logit"),
                        control = glmerControl(optimizer = "bobyqa"))

summary(model.grammatical.null)


anova(model.grammatical1, model.grammatical.null)

#POST-HOC for grammatical
post.hoc = emmeans(model.grammatical1, ~ proficiency)

pairs(post.hoc, adjust="tukey")

#####
#u.animate

#filter animate condition and NA answers
performance.u.animate = performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(condition == "u.animate")%>%
  filter(answer != "NA")%>%
  droplevels()
```

```
#table with relative frequency for the animate condition
performance.u.animate%>%
  group_by(proficiency, answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#animate graph
performance.u.animate%>%
  group_by(proficiency, answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = proficiency, y = frequency, fill = answer))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("cadetblue2", "lightpink3"))

model.u.animate = glm(as.factor(answer) ~ as.factor(proficiency), data = performance.u.animate,
family = binomial)
summary(model.u.animate)

#does not converge
model.u.animate1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                          (1+proficiency|item) + (1+proficiency|participant),
                        data = performance.u.animate, family = binomial(link = "logit"),
                        control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.u.animate1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                          (1+proficiency|item) + (1|participant),
                        data = performance.u.animate, family = binomial(link = "logit"),
                        control = glmerControl(optimizer = "bobyqa"))
#model 3 = works
model.u.animate1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                          (1|item) + (1|participant),
                        data = performance.u.animate, family = binomial(link = "logit"),
                        control = glmerControl(optimizer = "bobyqa"))

summary(model.u.animate1)

#model null
```

```
model.u.animate.null <- glmer(as.factor(answer)  ~ 1 +
                          (1|item) + (1|participant),
                        data = performance.u.animate, family = binomial(link = "logit"),
                        control = glmerControl(optimizer = "bobyqa"))

summary(model.u.animate.null)


anova(model.u.animate1, model.u.animate.null)

#POST-HOC for animate
post.hoc = emmeans(model.u.animate1, ~ proficiency)

pairs(post.hoc, adjust="tukey")

###
#u.inanimate

#filter inanimate condition and NA answers
performance.u.inanimate = performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(condition == "u.inanimate")%>%
  filter(answer != "NA")%>%
  droplevels()

#table with relative frequency for the inanimate condition
performance.u.inanimate%>%
  group_by(proficiency, answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#inanimate graph
performance.u.inanimate%>%
  group_by(proficiency, answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = proficiency, y = frequency, fill = answer))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("cadetblue2", "lightpink3"))
```

```
model.u.inanimate = glm(as.factor(answer) ~ as.factor(proficiency), data =
performance.u.inanimate, family = binomial)
summary(model.u.inanimate)

#does not converge
model.u.inanimate1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                    (1+proficiency|item) + (1+proficiency|participant),
                   data = performance.u.inanimate, family = binomial(link = "logit"),
                   control = glmerControl(optimizer = "bobyqa"))

##does not converge
model.u.inanimate1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                    (1+proficiency|item) + (1|participant),
                   data = performance.u.inanimate, family = binomial(link = "logit"),
                   control = glmerControl(optimizer = "bobyqa"))
#model 3 = works
model.u.inanimate1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                    (1|item) + (1|participant),
                   data = performance.u.inanimate, family = binomial(link = "logit"),
                   control = glmerControl(optimizer = "bobyqa"))

summary(model.u.inanimate1)

#model null
model.u.inanimate.null <- glmer(as.factor(answer)  ~ 1 +
                       (1|item) + (1|participant),
                      data = performance.u.inanimate, family = binomial(link = "logit"),
                      control = glmerControl(optimizer = "bobyqa"))

summary(model.u.inanimate.null)

anova(model.u.inanimate1, model.u.inanimate.null)

#POST-HOC for inanimate
post.hoc = emmeans(model.u.inanimate1, ~ proficiency)

pairs(post.hoc, adjust="tukey")
```

```
###
#u.pronoun

#filter pronoun condition and NA answers
performance.u.pronoun = performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(condition == "u.pronoun")%>%
  filter(answer != "NA")%>%
  droplevels()

#table with relative frequency for the pronoun condition
performance.u.pronoun%>%
  group_by(proficiency, answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#pronoun graph
performance.u.pronoun%>%
  group_by(proficiency, answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = proficiency, y = frequency, fill = answer))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("cadetblue2", "lightpink3"))


model.u.pronoun = glm(as.factor(answer) ~ as.factor(proficiency), data =
performance.u.pronoun, family = binomial)
summary(model.u.pronoun)

#does not converge
model.u.pronoun1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                          (1+proficiency|item) + (1+proficiency|participant),
                        data = performance.u.pronoun, family = binomial(link = "logit"),
                        control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.u.pronoun1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                          (1+proficiency|item) + (1|participant),
```

```
                    data = performance.u.pronoun, family = binomial(link = "logit"),
                    control = glmerControl(optimizer = "bobyqa"))
#model 3 = works
model.u.pronoun1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                      (1|item) + (1|participant),
                    data = performance.u.pronoun, family = binomial(link = "logit"),
                    control = glmerControl(optimizer = "bobyqa"))

summary(model.u.pronoun1)

#model null

model.u.pronoun.null <- glmer(as.factor(answer)  ~ 1 +
                         (1|item) + (1|participant),
                       data = performance.u.pronoun, family = binomial(link = "logit"),
                       control = glmerControl(optimizer = "bobyqa"))

summary(model.u.pronoun.null)

anova(model.u.pronoun1, model.u.pronoun.null )

#POST-HOC for pronoun
post.hoc = emmeans(model.u.pronoun1, ~ proficiency)

pairs(post.hoc, adjust="tukey")

###
#u.name

#filter name condition and NA answers
performance.u.name = performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(condition == "u.name")%>%
  filter(answer != "NA")%>%
  droplevels()

#table with relative frequency for the name condition
performance.u.name%>%
  group_by(proficiency, answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))
```

```
#name graph
performance.u.name%>%
  group_by(proficiency, answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = proficiency, y = frequency, fill = answer))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge")+
  theme_bw() +
  scale_fill_manual(values = c("cadetblue2", "lightpink3"))


model.u.name = glm(as.factor(answer) ~ as.factor(proficiency), data = performance.u.name,
family = binomial)
summary(model.u.name)

#does not converge
model.u.name1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                         (1+proficiency|item) + (1+proficiency|participant),
                       data = performance.u.name, family = binomial(link = "logit"),
                       control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.u.name1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                         (1+proficiency|item) + (1|participant),
                       data = performance.u.name, family = binomial(link = "logit"),
                       control = glmerControl(optimizer = "bobyqa"))
#model 3 = works
model.u.name1 <- glmer(as.factor(answer)  ~ as.factor(proficiency) +
                         (1|item) + (1|participant),
                       data = performance.u.name, family = binomial(link = "logit"),
                       control = glmerControl(optimizer = "bobyqa"))

summary(model.u.name1)

#model null

model.u.name.null <- glmer(as.factor(answer)  ~ 1 +
                             (1|item) + (1|participant),
                           data = performance.u.name, family = binomial(link = "logit"),
```

```
                     control = glmerControl(optimizer = "bobyqa"))

summary(model.u.name.null)

anova(model.u.name1, model.u.name.null )

#POST-HOC for name
post.hoc = emmeans(model.u.name1, ~ proficiency)

pairs(post.hoc, adjust="tukey")

############ INTERACTION################

#filter NA answers
performance.interaction= performance_groups_raters%>%
  select (condition, proficiency, item, participant, answer)%>%
  filter(answer != "NA")%>%
  droplevels()

model.interaction = glm(as.factor(answer) ~ as.factor(proficiency) * as.factor(condition),
                   data = performance.interaction, family = binomial)
summary(model.interaction)

#does not converge
model.interaction1 <- glmer(as.factor(answer) ~ as.factor(proficiency) * as.factor(condition) +
                   (1+proficiency*condition|item) + (1+proficiency*condition|participant),
                  data = performance.interaction, family = binomial(link = "logit"),
                  control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.interaction1 <- glmer(as.factor(answer) ~ as.factor(proficiency) * as.factor(condition) +
                     (1+proficiency*condition|item) + (1|participant),
                    data = performance.interaction, family = binomial(link = "logit"),
                    control = glmerControl(optimizer = "bobyqa"))

#converged
model.interaction1 <- glmer(as.factor(answer) ~ as.factor(proficiency) * as.factor(condition) +
                     (1|item) + (1|participant),
                    data = performance.interaction, family = binomial(link = "logit"),
                    control = glmerControl(optimizer = "bobyqa"))
```

```
summary(model.interaction1)

#model without interaction

model.no.inter <- glmer(as.factor(answer) ~ as.factor(proficiency) + as.factor(condition) +
                          (1|item) + (1|participant),
                        data = performance.interaction, family = binomial(link = "logit"),
                        control = glmerControl(optimizer = "bobyqa"))

summary(model.no.inter)

anova(model.interaction1, model.no.inter)

#POST-HOC for interaction

post.hoc = emmeans(model.interaction1, ~ proficiency*condition)

pairs(post.hoc, adjust="tukey")

###################################IMAGES###################################

library(readr)
performance_images_combined <- read_csv("performance.images.combined.csv")
View(performance_images_combined)

#Inspect the first lines
head(performance_images_combined)

#Check the variables
str(performance_images_combined)

#Let's check the values at "images"
unique(performance_images_combined$images_answer)
#Let's check the values at "proficiency"
unique(performance_images_combined$proficiency)
#Let's check the values at "item"
unique(performance_images_combined$item)
#Let's check the values at "participant"
unique(performance_images_combined$participant)

#Expected answer  = correct answer
```

```
#Unexpected answer = incorrect answer
#NA = missed answer

#filter NA answers in the column images
performance.images = performance_images_combined%>%
  select(participant, images_answer, proficiency, item)%>%
  filter(images_answer != "NA")%>%
  droplevels()

#table with the relative frequency of expected/unexpected answers in each proficiency group
performance.images%>%
  group_by(proficiency, images_answer)%>%
  summarise(n = n())%>%
  mutate(freq = n / sum(n))

#images graph
performance.images%>%
  group_by(proficiency, images_answer)%>%
  summarise(n = n ())%>%
  mutate(frequency = n/sum(n))%>%
  ggplot(., aes(x = proficiency, y = frequency, fill = images_answer))+
  scale_y_continuous(labels = scales::percent)+
  geom_col(position = "dodge") +
  theme_bw() +
  scale_fill_manual(values = c("steelblue1", "tan1"))

#model
model.images = glm(as.factor(images_answer) ~ as.factor(proficiency), data =
performance.images, family = binomial)

summary(model.images)

#does not converge
model.images1 <- glmer(as.factor(images_answer)  ~ as.factor(proficiency) +
                         (1+proficiency|item) + (1+proficiency|participant),
                       data = performance.images, family = binomial(link = "logit"),
                       control = glmerControl(optimizer = "bobyqa"))

#does not converge
model.images1 = glmer(as.factor(images_answer)  ~ as.factor(proficiency) +
                      (1+proficiency|item) + (1|participant),
```

```
                 data = performance.images, family = binomial(link = "logit"),
                 control = glmerControl(optimizer = "bobyqa"))

#model 3 = works
model.images1 = glmer(as.factor(images_answer)  ~ as.factor(proficiency) +
                   (1|item) + (1|participant),
                 data = performance.images, family = binomial(link = "logit"),
                 control = glmerControl(optimizer = "bobyqa"))


summary(model.images1)

#model null
model.images.null = glmer(as.factor(images_answer)  ~ 1 + (1|item) + (1|participant),
  data = performance.images, family = binomial(link = "logit"),
  control = glmerControl(optimizer = "bobyqa"))

summary(model.images.null)

anova(model.images1,model.images.null)

###POST-HOC for images

post.hoc = emmeans(model.images1, ~ proficiency)

pairs(post.hoc, adjust="tukey")
```

Appendix 9

Table 9: Contrast between Conditions and Proficiency Groups

contrast	estimate	SE	z.ratio	p.value
higher proficiency grammatical - lower proficiency grammatical	-3.1290	1.116	-2.805	0.2438
higher proficiency grammatical - native grammatical	3.5953	1.891	1.901	0.8530
higher proficiency grammatical - higher proficiency u.animate	-3.0360	0.637	-4.762	0.0002*
higher proficiency grammatical - lower proficiency u.animate	-6.4067	1.122	-5.708	<0.0001*
higher proficiency grammatical - native u.animate	-3.3696	1.336	-2.522	0.4241
higher proficiency grammatical - higher proficiency u.inanimate	-3.3558	0.639	-5.248	<0.0001*
higher proficiency grammatical - lower proficiency u.inanimate	-6.6973	1.127	-5.945	<0.0001*
higher proficiency grammatical - native u.inanimate	-3.1461	1.337	-2.354	0.5499
higher proficiency grammatical - higher proficiency u.name	-1.8780	0.620	-3.031	0.1413
higher proficiency grammatical - lower proficiency u.name	-6.1724	1.118	-5.521	<0.0001*
higher proficiency grammatical - native u.name	-3.5947	1.335	-2.693	0.3088
higher proficiency grammatical - higher proficiency u.pronoun	-2.0523	0.620	-3.310	0.0642
higher proficiency grammatical - lower proficiency u.pronoun	-5.9256	1.116	-5.310	<0.0001*

higher proficiency grammatical - native u.pronoun	-4.0444	1.331	-3.039	0.1387
lower proficiency grammatical - native grammatical	6.7243	1.844	3.647	0.0213*
lower proficiency grammatical - higher proficiency u.animate	0.0930	1.038	0.090	1.0000
lower proficiency grammatical - lower proficiency u.animate	-3.2778	0.432	-7.590	<0.0001*
lower proficiency grammatical - native u.animate	-0.2406	1.238	-0.194	1.0000
lower proficiency grammatical - higher proficiency u.inanimate	-0.2268	1.030	-0.220	1.0000
lower proficiency grammatical - lower proficiency u.inanimate	-3.5683	0.445	-8.012	<0.0001*
lower proficiency grammatical - native u.inanimate	-0.0171	1.240	-0.014	1.0000
lower proficiency grammatical - higher proficiency u.name	1.2510	1.069	1.170	0.9980
lower proficiency grammatical - lower proficiency u.name	-3.0435	0.417	-7.302	<0.0001*
lower proficiency grammatical - native u.name	-0.4657	1.235	-0.377	1.0000
lower proficiency grammatical - higher proficiency u.pronoun	1.0767	1.064	1.012	0.9996
lower proficiency grammatical - lower proficiency u.pronoun	-2.7967	0.409	-6.844	<0.0001*
lower proficiency grammatical - native u.pronoun	-0.9154	1.229	-0.745	1.0000
native grammatical - higher proficiency u.animate	-6.6313	1.860	-3.565	0.0282*

native grammatical - lower proficiency u.animate	-10.0020	1.854	-5.395	<0.0001*
native grammatical - native u.animate	-6.9648	1.584	-4.397	0.0010*
native grammatical - higher proficiency u.inanimate	-6.9511	1.857	-3.743	0.0151*
native grammatical - lower proficiency u.inanimate	-10.2926	1.857	-5.543	<0.0001*
native grammatical - native u.inanimate	-6.7414	1.576	-4.278	0.0018*
native grammatical - higher proficiency u.name	-5.4733	1.871	-2.925	0.1847
native grammatical - lower proficiency u.name	-9.7677	1.851	-5.278	<0.0001*
native grammatical - native u.name	-7.1899	1.591	-4.518	0.0006*
native grammatical - higher proficiency u.pronoun	-5.6476	1.869	-3.022	0.1449
native grammatical - lower proficiency u.pronoun	-9.5209	1.849	-5.149	<0.0001*
native grammatical - native u.pronoun	-7.6397	1.604	-4.762	0.0002*
higher proficiency u.animate - lower proficiency u.animate	-3.3708	1.042	-3.235	0.0805
higher proficiency u.animate - native u.animate	-0.3336	1.276	-0.261	1.0000
higher proficiency u.animate - higher proficiency u.inanimate	-0.3198	0.476	-0.672	1.0000
higher proficiency u.animate - lower proficiency u.inanimate	-3.6613	1.046	-3.499	0.0352*
higher proficiency u.animate -	-0.1101	1.277	-0.086	1.0000

native u.inanimate				
higher proficiency u.animate - higher proficiency u.name	1.1580	0.548	2.111	0.7275
higher proficiency u.animate - lower proficiency u.name	-3.1364	1.037	-3.023	0.1443
higher proficiency u.animate - native u.name	-0.5587	1.274	-0.438	1.0000
higher proficiency u.animate - higher proficiency u.pronoun	0.9837	0.538	1.829	0.8868
higher proficiency u.animate - lower proficiency u.pronoun	-2.8897	1.036	-2.791	0.2517
higher proficiency u.animate - native u.pronoun	-1.0084	1.269	-0.794	1.0000
lower proficiency u.animate - native u.animate	3.0372	1.242	2.446	0.4802
lower proficiency u.animate - higher proficiency u.inanimate	3.0509	1.034	2.950	0.1735
lower proficiency u.animate - lower proficiency u.inanimate	-0.2905	0.386	-0.752	1.0000
lower proficiency u.animate - native u.inanimate	3.2607	1.244	2.621	0.3553
lower proficiency u.animate - higher proficiency u.name	4.5288	1.075	4.213	0.0023*
lower proficiency u.animate - lower proficiency u.name	0.2343	0.368	0.637	1.0000
lower proficiency u.animate - native u.name	2.8121	1.239	2.270	0.6134
lower proficiency u.animate - higher proficiency u.pronoun	4.3545	1.069	4.073	0.0042*
lower proficiency u.animate - lower proficiency u.pronoun	0.4811	0.366	1.313	0.9933

lower proficiency u.animate - native u.pronoun	2.3624	1.232	1.918	0.8448
native u.animate - higher proficiency u.inanimate	0.0137	1.270	0.011	1.0000
native u.animate - lower proficiency u.inanimate	-3.3277	1.245	-2.672	0.3220
native u.animate - native u.inanimate	0.2235	0.669	0.334	1.0000
native u.animate - higher proficiency u.name	1.4916	1.300	1.147	0.9983
native u.animate - lower proficiency u.name	-2.8029	1.238	-2.264	0.6174
native u.animate - native u.name	-0.2251	0.672	-0.335	1.0000
native u.animate - higher proficiency u.pronoun	1.3173	1.296	1.017	0.9996
native u.animate - lower proficiency u.pronoun	-2.5561	1.236	-2.068	0.7567
native u.animate - native u.pronoun	-0.6748	0.677	-0.997	0.9997
higher proficiency u.inanimate - lower proficiency u.inanimate	-3.3414	1.038	-3.218	0.0845
higher proficiency u.inanimate - native u.inanimate	0.2098	1.272	0.165	1.0000
higher proficiency u.inanimate - higher proficiency u.name	1.4779	0.543	2.719	0.2928
higher proficiency u.inanimate - lower proficiency u.name	-2.8166	1.029	-2.736	0.2828
higher proficiency u.inanimate - native u.name	-0.2388	1.268	-0.188	1.0000
higher proficiency u.inanimate - higher proficiency u.pronoun	1.3035	0.532	2.452	0.4759
higher proficiency u.inanimate -	-2.5698	1.028	-2.501	0.4396

lower proficiency u.pronoun				
higher proficiency u.inanimate - native u.pronoun	-0.6885	1.263	-0.545	1.0000
lower proficiency u.inanimate - native u.inanimate	3.5512	1.248	2.846	0.2221
lower proficiency u.inanimate - higher proficiency u.name	4.8193	1.079	4.466	0.0008*
lower proficiency u.inanimate - lower proficiency u.name	0.5248	0.378	1.387	0.9886
lower proficiency u.inanimate - native u.name	3.1026	1.243	2.497	0.4424
lower proficiency u.inanimate - higher proficiency u.pronoun	4.6450	1.073	4.327	0.0014*
lower proficiency u.inanimate - lower proficiency u.pronoun	0.7716	0.377	2.044	0.7718
lower proficiency u.inanimate - native u.pronoun	2.6529	1.235	2.147	0.7028
native u.inanimate - higher proficiency u.name	1.2681	1.301	0.975	0.9997
native u.inanimate - lower proficiency u.name	-3.0264	1.240	-2.440	0.4842
native u.inanimate - native u.name	-0.4486	0.672	0.667	1.0000
native u.inanimate - higher proficiency u.pronoun	1.0938	1.297	0.843	1.0000
native u.inanimate - lower proficiency u.pronoun	-2.7796	1.238	-2.244	0.6322
native u.inanimate - native u.pronoun	-0.8983	0.681	-1.319	0.9930
higher proficiency u.name - lower proficiency u.name	-4.2945	1.070	-4.012	0.0054*

higher proficiency u.name - native u.name	-1.7167	1.298	-1.322	0.9928
higher proficiency u.name - higher proficiency u.pronoun	-0.1743	0.561	-0.311	1.0000
higher proficiency u.name - lower proficiency u.pronoun	-4.0477	1.068	-3.789	0.0127*
higher proficiency u.name - native u.pronoun	-2.1664	1.294	-1.674	0.9411
lower proficiency u.name - native u.name	2.5778	1.235	2.087	0.7439
lower proficiency u.name - higher proficiency u.pronoun	4.1202	1.065	3.870	0.0094*
lower proficiency u.name - lower proficiency u.pronoun	0.2468	0.356	0.694	1.0000
lower proficiency u.name - native u.pronoun	2.1281	1.228	1.733	0.9233
native u.name - higher proficiency u.pronoun	1.5424	1.294	1.192	0.9975
native u.name - lower proficiency u.pronoun	-2.3310	1.233	-1.890	0.8588
native u.name - native u.pronoun	-0.4497	0.673	-0.668	1.0000
higher proficiency u.pronoun - lower proficiency u.pronoun	-3.8734	1.063	-3.645	0.0214*
higher proficiency u.pronoun - native u.pronoun	-1.9921	1.290	-1.545	0.9696
lower proficiency u.pronoun - native u.pronoun	1.8813	1.227	1.534	0.9714

* Statistically significant difference