University of Reading

Non-Native Sentence Processing: A Focus on Agreement and Reflexives

A thesis submitted in fulfilment of requirements for the degree of Doctor of Philosophy

Department of English Language and Applied Linguistics

Shatha Alaskar

September 2023

Abstract

An ongoing debate revolves around whether non-native (L2) speakers resolve discontinuous linguistic dependencies, such as subject-verb agreement and reflexive dependencies, similarly to native (L1) speakers. Variable findings across studies have contributed to this ongoing discussion. Among the different factors suggested to explain potential L1/L2 processing differences, is the role that the working memory retrieval mechanism, which subserves dependency processing, may play in L2 processing. While effects associated with this retrieval mechanism, namely similarity-based interference effects, have been extensively investigated in L1 processing, their relevance to L2 processing remains speculative.

The question of whether the computation of subject-verb agreement and reflexive dependencies relies on the same memory retrieval mechanism has also been a subject of debate in the L1 processing literature. This thesis aims to examine these two dependencies, subject-verb agreement and reflexive-antecedent dependencies, in parallel experimental settings within the same group of participants at different proficiency levels. This approach offers a valuable means to understand L2 processing patterns across dependencies and brings us closer to reaching a general agreement on L1/L2 processing similarities and differences, while minimizing individual variation.

Through two large-scale studies employing offline and online measures, the results generally indicated that L2 speakers' processing patterns are largely similar to those of L1 speakers. L2 speakers demonstarted sensitivity to grammatical constraints without an increased vulnerability to interference effects compared to L1 speakers. Reduced grammaticality effects, however, were observed specifically for S-V agreement in L2 speakers, while no such attenuation was found for reflexive-antecedent dependencies when compared to L1 speakers. Proficiency

analysis revealed that this discrepancy may be driven by the performance of lower proficiency L2 speakers in S-V agreement, contradicting the more target-like processing profile observed in higher proficiency L2 speakers.

Taken together, the overall pattern of results suggests that L1/L2 processing follows a similar manner, and any potential quantitative differences can be attributed to individual differences, particularly in L2 proficiency.

Acknowledgements

My heartfelt gratitude, first and foremost, goes to my supervisor Dr Ian Cunnings, whose unwavering support and guidance were instrumental in the successful completion of this thesis. I sincerely appreciate his generosity in sharing his knowledge, and the invaluable feedback he provided on conference abstracts/presentations, paper manuscripts, and my thesis. His constant dedication and availability have been a source of strength, comfort, and encouragement throughout my journey. I am also grateful for the opportunities he offered me to work on side projects, which allowed me to learn about eye-tracking data collection and analysis. Indeed, he played a crucial role in making my PhD journey more enjoyable and rewarding than I ever thought possible, and I couldn't have asked for a better supervisor. My gratitude additionally to Dr Fraibet Aveledo for her support, and for my examiners, Prof. Leah Roberts and Prof. Ludovica Serratrice, for their thought-provoking discussion and valuable feedback.

I am also immensely grateful to all the participants who displayed benevolence by generously devoting their time. Without their participation, these studies would not have been possible. I extend my appreciation to Majmaah University and the Saudi Cultural Bureau for their assistance and financial support, which enabled me to pursue my academic and professional aspirations. I would also like to thank the DELAL department, especially Prof. Parvaneh Tavakoli, and the Psycholinguistics and Neurolinguistics Lab at the School of PCLS, which I am truly glad to have joined, for their kind support. Thanks also to my friends for their encouragement and uplifting words during these years.

Last but certainly not least, I owe undying gratitude to my family, who have been the pillars of my strength and support throughout my journey. To my parents, I am incredibly blessed to have you by my side, always believing in me and encouraging me to pursue my aspirations. To

my siblings Ahmed, Abeer, Areej and Yasmeen, thank you for your endless support. You have always been there to lift my spirits, offer a helping hand and show genuine care in countless ways. To my beloved grandmother, aunts and uncles, thank you for your care and sincere wishes. Special thanks to my aunt, Dr Asma, my rock and lifelong mentor since the day I was born. To all, your love and support will always be cherished.

Declaration of Authorship

I confirm that this is my own work and the use of all materials from other sources has been properly and fully acknowledged.

Shatha Alaskar

Table of Contents

List of Tables

List of Figures

CHAPTER ONE

1. Introduction

Forming syntactic dependencies between non-adjacent constituents is a prerequisite for successful language comprehension. For instance, matching a verb with its grammatical controller, the subject, as in (1), or linking anaphoric expressions, such as reflexives, to their antecedents as in (2), correctly and rapidly requires the integration of different sources of information in real time. This process provides the parser with information about structural relations in the context, which ultimately contributes to successful comprehension.

(1) The key to the cabinet was rusty.

(2) The boy near the man hurt himself.

The processing of such dependencies has informed debate about native (L1) and non-native (L2) processing, but how to explain potential L1/L2 differences is contested. An influential account posits that L2 speakers, defined in this thesis as those who started learning an L2 after acquiring their native language at the earliest after age 5 compared to native speakers who have learnt the language since birth, have difficulty applying syntactic constraints during processing relative to L1 speakers, which may create different parsing mechanisms (The Shallow Structure Hypothesis, SHH, Clahsen & Felser, 2006a, 2018). Others, alternatively, attribute L1/L2 potential differences to cognitive demands or individual differences in proficiency or lexical processing ability, rather than differences in syntactic processing per se (Hopp, 2014; Lim & Christianson, 2014; McDonald, 2006; Sagarra & Herschensohn, 2010). More recently, Cunnings (2017) argued that L2 speakers can employ the same syntactic parses as L1 speakers, and differences can be rather explained in terms of working memory operations that underpin sentence comprehension.

Against this background, whether L2 speakers violate constraints on linguistic dependencies during processing has been important in assessing these theories, but existing research has revealed mixed findings. While some studies suggest L1 speakers and L2 speakers compute linguistic dependencies in a similar manner, others indicate that they resolve dependencies differently (Dallas & Kaan, 2008; Felser, 2015; see Roberts, 2013 for a review). A number of factors may influence these findings. L2 speakers' performance may vary depending on the type of dependency, the L1-L2 combinations, the measures used and the participants' individual characteristics. Therefore, we contribute to this theoretical debate by examining subject-verb (S-V) agreement and reflexive-antecedent dependencies. We are unaware of any existing study that has examined these two dependencies in the same L2 group. In the L1 processing literature, whether S-V agreement and reflexive resolution are resolved via the same mechanism during sentence processing is debated (e.g., Dillon, Mishler, Sloggett, & Phillips, 2013; Jäger, Mertzen, Van Dyke, & Vasishth, 2020). In light of this variation, this thesis compares whether the same mechanism is employed when resolving these two dependencies in L2 processing, as well as whether L1 and L2 processing differ in this respect.

Given that S-V agreement and reflexive dependencies share similar properties in Arabic and English (see Chapter 2.2.2), these linguistic phenomena were chosen so that L1 influence that results from differences between languages can be diminished and complete focus will be given to the mechanisms that underpin L2 processing. Apart from the similarity of S-V agreement and reflexives in both languages, both dependencies require the retrieval of a subject during incremental processing for dependency formation, and they both require the subject to generally show the same agreement features as the dependent element (the verb or reflexive) that similarly distinguishes different feature values. In this sense, there is an agreement constraint or relation between the dependent elements, although agreement for S-V agreement is based on morphosyntactic indexing for event participants (Corbett, 2006), while reflexive-antecedent

agreement is part of a referential relation or dependency between two noun phrases that is primarily based on binding condition A (Baker, 2008; Chomsky, 1981; Dillon et al., 2013; Kratzer, 2009).

Processing such linguistic relations have been argued to rely on a cue-based memory retrieval mechanism that deploys the linguistic cues associated with a given dependency from linguistic memory to resolve that dependency (Lewis & Vasishth, 2005; Lewis, Vasishth, & Dyke, 2006; McElree, 2000; McElree, Foraker, & Dyer, 2003). By way of illustration, in the following classic example taken from Pearlmutter, Garnsey and Bock (1999), the main verb 'was' needs to be connected to its target subject 'the key' to complete the S-V agreement dependency.

(3) a. **The key** $\begin{pmatrix} +\,subject \\ +\,singular \end{pmatrix}$ to the cabinet $(+\,singular)$ **was** $\begin{Bmatrix} +\,subject \\ +singular \end{Bmatrix}$ rusty …
b. **The key** $\begin{pmatrix} +\,subject \\ +\,singular \end{pmatrix}$ to the cabinets $(+\,plural)$ **was** $\begin{Bmatrix} +\,subject \\ +singular \end{Bmatrix}$ rusty …
c. ***The key** $\begin{pmatrix} +\,subject \\ +\,singular \end{pmatrix}$ to the cabinets $(+\,plural)$ **were** $\begin{Bmatrix} +\,subject \\ +plural \end{Bmatrix}$ rusty …
d. ***The key** $\begin{pmatrix} +\,subject \\ +\,singular \end{pmatrix}$ to the cabinet $(+\,singular)$ **were** $\begin{Bmatrix} +\,subject \\ +plural \end{Bmatrix}$ rusty …

The reflexive-antecedent dependency in example (4) also needs to be resolved to comprehend the sentence correctly. In order to achieve this, the reflexive 'himself/herself' should be bound by a c-commanding subject within the local domain, here, 'the schoolboy' as per Principle A of Binding Theory (Chomsky, 1981, 1986).

(4) a. **The schoolboy** $\begin{pmatrix} +\,subject \\ +\,c-command \\ +\,masculine \end{pmatrix}$ who asked the man $(+\,masculine)$ had prepared **himself** $\begin{Bmatrix} +\,subject \\ +\,c-command \\ +\,masculine \end{Bmatrix}$ …

b. **The schoolboy** $\begin{pmatrix} +\,subject \\ +\,c-command \\ +\,masculine \end{pmatrix}$ who asked the woman $(+\,feminine)$ had prepared **himself** $\begin{Bmatrix} +\,subject \\ +\,c-command \\ +\,masculine \end{Bmatrix}$ …

c. **The schoolboy** $\begin{pmatrix} +\,subject \\ +\,c-command \\ +\,masculine \end{pmatrix}$ who asked the woman $(+\,feminine)$ had prepared **herself** $\begin{Bmatrix} +\,subject \\ +\,c-command \\ +feminine \end{Bmatrix}$ …

d. **The schoolboy** $\begin{pmatrix} +subject \\ +c-command \\ +masculine \end{pmatrix}$ who asked the man $(+masculine)$ had prepared **herself** $\begin{Bmatrix} +subject \\ +c-command \\ +feminin \end{Bmatrix} \dots$

According to cue-based parsing, the parser accesses memory representations when reaching the verb in (3) or reflexive in (4) to retrieve the target item ('the key' in (3) or 'the schoolboy' in (4)) for dependency resolution. Retrieval cues derived from the local syntactic context and the item that triggered this process will be used to guide retrieval. For instance, $\{+subject\}$ and $\{+singular\}$ are set out by the verb in (3a/b) to seek out a matching noun that can act as the subject of the verb 'was'. Similarly, the reflexive in (4a/b) may use $\{+subject\}$, $\{+c-command\}$[1] and $\{+masculine\}$ to find an antecedent with matching features. However, the presence of similar items or 'distractors' in memory that partially match the retrieval cues may decrease the possibility of retrieving the target item and potentially cause retrieval errors. This is known as similarity-based interference (Lewis et al., 2006; Nicenboim & Vasishth, 2018; Parker, Shvartsman, & Van Dyke, 2017; Vasishth, Nicenboim, Engelmann, & Burchert, 2019). For example, in (3a/c), the complex noun phrase headed by 'the key' also includes a distractor ('the cabinet/s') that may interfere during dependency resolution when it matches the verb number. The overlap in gender between the distractor ('man/woman') and the reflexive in (4a/c) can also increase the possibility of wrongly retrieving the partially matching distractor. This similarity-based interference is predicted to possibly cause different effects on readers' processing patterns (see Chapter 2.2.3 for further detailed discussion).

Some L1 processing studies have argued though that susceptibility to interference effects during S-V agreement and reflexive-antecedent dependencies processing may differ, such that S-

[1] Technically, c-command is a relational syntactic concept that describes the relation between two constituents, rather than a static feature. Therefore, a more complex computation is involved in tracking c-command relative to static features (Jäger, Engelmann, & Vasishth, 2017). This distinction is beyond the scope of the thesis, but see Kush (2013) for further discussion.

V agreement is prone to interference effects whereas reflexives resist interference effects (Dillon et al., 2013). This has led to the suggestion that the reflexive dependency completion process is guided only by syntactic cues, while agreement features that can be obtained from lexical items during the retrieval process are not used as is the case for S-V agreement (Dillon et al., 2013). Others, however, argued that both dependencies can similarly be susceptible to similarity-based interference effects (Jäger et al., 2020). As how readers resolve these two dependencies has been debated in L1 processing, we also aim to assess this controversy in L2 processing.

Inspired by this increasing interest in memory mechanisms that subserve sentence processing, Cunnings (2017) proposed a new perspective on L2 processing arguing that L2 speakers' success in resolving a dependency might depend on the properties of the other constituents in the sentence. That is, the similarity of other constituents to the target dependent element that needs to be retrieved may influence the potentiality to locate the target element in memory and retrieve it. With respect to L1/L2 processing differences, L2 speakers were specifically argued to be more susceptible to similarity-based interference effects than L1 speakers. A highlight of the thesis thus is its focus on a rarely explored aspect of L2 processing and which has recently been proposed by Cunnings (2017), similarity-based interference during cue-based memory retrieval parsing. To avoid terminological confusion, it is worth notifying here that 'L1 influence', 'L1 transfer' and especially 'L1 interference' are notions found and studied independently in second language acquisition literature (SLA) to refer to the influence of one language on the other and should not be confused with the term 'similarity-based interference'. The term 'interference', as a shorthand to 'similarity-based interference', that will be used throughout this thesis is discussed differently in sentence processing literature and will only be used to refer to the potential outcome that may result from processing similar constituents within a sentence in line with the similarity-based interference phenomenon. As such, 'interference' as discussed within this thesis relates to interference from constituents within a sentence during

processing, while I will use the term ‘transfer’ to describe the potential influence of one language on another.

Taken together, in two large-scale studies using offline and online measures, our goal is to examine the processing patterns of L2 speakers during dependency resolution in S-V agreement and reflexives. We aim to explore whether working memory operations, specifically the memory retrieval mechanism, can account for any potential L1/L2 processing differences. Additionally, we will examine whether alternative explanations of L2 processing can clarify our findings. This thesis is structured as follows. The next chapter begins by laying out the linguistic background of S-V agreement and reflexive dependencies in English and Arabic, followed by a thorough review of relevant theoretical dimensions and empirical literature related to the linguistic phenomena under study. It will then proceed to discuss research aims and predictions. Chapter 3 presents Study 1 that examined S-V agreement and reflexive-antecedent dependencies using grammaticality judgement and self-paced reading tasks along with other individual differences measures. To further evaluate the findings of Study 1, Chapter 4 presents Study 2 which examined reflexive-antecedent dependencies but with a number manipulation rather than gender as was done in Study 1. Chapter 5, finally, concludes this thesis with a general discussion of the findings considering different theoretical accounts on factors that may influence L2 processing and assessing which can best explain L2 speakers’ performance. To briefly summarise the main findings from these experiments, the results suggest that L1 and L2 speakers resolve dependencies similarly, but individual differences in proficiency appear to modulate L2 processing patterns.

CHAPTER TWO

2. Literature Review

2.1 Overview

This chapter initially describes the linguistic context of S-V agreement and reflexives in English and Arabic, followed by a discussion of L1/L2 theoretical accounts to sentence processing, with a specific focus on working memory conceptualisations and cue-based memory retrieval interference. Empirical research relevant to the focus of the study will be reported from L1 and L2 literature. The chapter then concludes with a detailed description of the research aims and questions that will be addressed.

2.2 Linguistic Background

2.2.1 Subject-verb agreement dependency.

Online grammatical integration of structural features manifested morphologically on individual lexical elements is an essential aspect of sentence comprehension in real time. This is evident in the case of subject-verb agreement, a relationship displayed in many languages in which verbs generally change in form according to subject features. Subject-verb agreement dependency is the first linguistic phenomenon considered in this research to investigate L2 processing. As languages vary in the degree to which their morphosyntactic systems demonstrate conceptual and formal distinctions, subject-verb agreement in English and Arabic will be discussed below.

2.2.1.1 Subject-Verb Agreement in English.

English has little agreement morphology, but subject-verb agreement is one place where it is found. Agreement in number is the most fundamental type of concord considered in agreement studies. Number feature is reflected overtly in English on nominals and most verbs. The plural morphological inflection ‘-s’ that number has on nouns means numerous entities in the

real world and its absence, generally, indicates one entity. With regards to verbs, English has two compositionally different verbs: thematic verbs that can be divided into smaller morphological units as a stem and bound morpheme (e.g., write+*s*) and copulas or verbs *be* (am, is, are, was and were). Number manifestation on verbs is syntactically dependent and can be clearly demonstrated by the copulas and the present 3rd person singular verb that is inflected with -*s* among other forms[2] as illustrated below.

(5) **Thematic Verb**

a. The student writes a letter every day.

b. The students write a letter every day.

(6) **Copula (Verb *be*)**

a. The boy is a student in this school.

b. The boys are students in this school.

In (5a/6a), the subjects 'the student/the boy' are singular noun phrases (NPs) with 'zero' or 'Ø' singular number morphology, which signifies a single real-world entity and hence requires a singular verb. Such verbs can be a thematic verb as in (5a) in which the verb 'write*s*' should be inflected with -*s* that denotes present 3rd person singular to agree with the singular subject or a copula *be* as in (6a) where the verb (is) is an individual lexical item that specifies the singular number feature per se as a free morpheme. Likewise, the plurality of subjects 'the students/boys' in (5b/6b) is displayed by the attached bound morpheme -*s* signifying that a plural verb is needed. If the context requires a copula *be*, such as in (6b), the appropriate verb (in this case, 'are') would

[2] The past tense of copulas also constitutes number, e.g. He *was* and we *were* whereas the past tense of thematic verbs does not. However, tense inflection is not within the scope of this research.

be sought to agree with the plural subject. However, in the case of a thematic verb, there would be no overt morphology marking. Infl and related agreement features remain evident, however. In the example, ('the students write…'), the verb contains features for person (3rd) and number (plural), which happen to be not explicitly manifested. Although overt verbal suffixes do not demonstrate these abstract features, we realise that they should be present at some level due to various agreement patterns (White, 2003, p. 180). Thus, replacing the singular verb in (5a/6a) with a plural verb or vice versa for (5b/6b) leads to ungrammaticality because of the mismatch between the verb and its subject in number.

2.2.1.2 Subject-Verb Agreement in Arabic.

In Modern Standard Arabic (henceforth, MSA or just 'Arabic'), number is also taken into account when a subject and a verb enter into an agreement relation. As for agreement morphology on nominals, number feature is realised morphologically and is significant for the process of S-V agreement. Similar to English, unmarked nouns are generally singular, but dual and plural nominals have number features morphologically marked, as demonstrated in Table 1[3]. Verbs additionally mark morphologically for number in the same way as nominals as shown in Table 2[4].

Table 1

Nouns in MSA in which morphological endings denote number values.

Masculine	Feminine	Translation
muʕallim	muʕallimah	One teacher (singular)
muʕallim-a:n	muʕallim-ata:n	Two teachers (dual)
muʕallim-u:n	muʕallim-a:t	More than two teachers (plural)

[3] The morphological suffixes used here are for the nominative case.
[4] The verbs in the table are imperfective in the present, but S-V number agreement is obligatory regardless of the tense.

Table 2

Thematic verbs in MSA in which morphological affixes denote number values.

Masculine	Feminine	Translation
ya-drusu	ta-drusu	He/She studies (singular)
ya-drus-a:n	ta-drus-a:n	They study (dual)
ya-drus-u:n	ta-drus-na	They study (plural)

Having described agreement morphology for nominals and verbs, the following part will discuss agreement patterns based on the sentence type. The subject in MSA can be either preverbal or postverbal, unlike English where the subject is only preverbal (SVO). Thus, MSA has two word orders (SVO and VSO), and agreement morphosyntactic features interact with word order. That is, the verb agrees fully with the subject in gender, person and number when it is singular regardless of word order. However, the case is different for dual and plural nouns. If the subject precedes the verb (SVO), the verb shows full agreement agreeing with the subject in all features including person, gender, and number. However, if the subject follows the verb (VSO), the verb matches the subject in gender and person but not in number as the verb here should be singular[5] (Mohammad, 2000). Generally speaking, both word orders are similarly common and used alternatively in MSA (Benmamoun, 1992; Fehri, 1993; Mohammad, 1989), and (SVO) word order is mostly the focus of agreement studies (Benduhaish, 2018).

Since subject-verb agreement in English has been reviewed in terms of thematic and copula verbs, English copulas' counterparts in MSA will also be discussed. While English has only one sentence type that contains a verb in its structure, Arabic has two sentence types: the verbal sentence and the so-called nominal (equational) sentence. The verbal sentence is a sentence

[5] The discussion here is about agreement in MSA, but the verb in many local varieties of Arabic, including Saudi Arabic, shows full agreement with dual and plural subjects regardless of the subject position (Musabhien, 2008).

that contains a thematic verb while the nominal (equational) sentence has no thematic verb but includes a subject and predicate. The latter type of sentence is equivalent to English sentences in which the main verb is the verb *be* in the present tense (Tahir, 2009). Arabic copulas, however, are expressed according to tense and context. In the present tense, the copula does not have a lexical description, but it is understood from the context. However, it appears explicitly in past contexts in the form of 'ka:na/ka:nu' that corresponds to 'was/were' as shown in (7).

(7) **almuʕallimu** alwaqıf khalf altˁalebi **ka:na** fi: alfasˁl

The teacher-3M.SG behind the student-3M.SG be-Prf.3M.SG in the class

"**The teacher**$_j$ behind the student$_i$ **was**$_{i*/j}$ in the class"

In the above sentence, there are two nominals 'almuʕallimu/the teacher' which is the target subject of the verb 'ka:na/was' and 'altˁalebi/the student' which is a noun modifying the subject. The verb 'ka:na/was' should solely be matched with its grammatical controller (the subject) 'almuʕallimu' disregarding any other nearby but structurally illicit nominals as 'altˁalebi' for successful sentence comprehension. As in English, replacing the verb 'ka:na' with 'ka:nu:', which is the plural form of the verb, would result in an ungrammatical sentence (e.g., '*a;muʕallimu alwaqıf khalf altˁalebi ka:nu: fi: alfasˁl.') because the verb does not match the target subject in number.

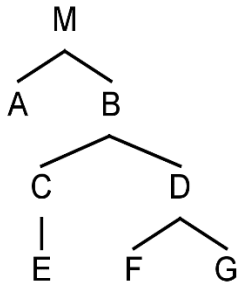
2.2.2 Reflexive-antecedent dependency.

To interpret sentences and context effectively, it is crucial to establish syntactic dependencies between non-adjacent elements. This is exemplified in the case of reflexive resolution, where the reader must connect the reflexive anaphor (such as 'himself' or 'herself') to the antecedent noun phrase upon which it depends referentially. Reflexive-antecedent

dependency is the second syntactic phenomenon used in the current endeavour to investigate L2 processing.

2.2.2.1 Reflexives and Binding Principle A.

Reflexive binding has conventionally been accounted for by the Government and Binding Theory (Chomsky, 1981, 1986). The Binding Principles are concerned with the relations between different constituents in a sentence, as in the case, for example, with anaphoric expressions, and it was defined by Chomsky as a set of constraints on the co-occurrence of nominal elements in a particular context. According to Principle A of the Binding Theory that was formulated to explain the restriction on reflexives interpretation, reflexive anaphors (e.g., herself, himself, etc.) must be bound by the closest c-commanding antecedent within the same local domain, namely its binding category. The concept of 'c-command' refers to the relationship between nodes in a grammatical parse tree based on the notion of dominance: node A dominates node C if node A is above node C in the syntactic tree. (Reinhart, 1983). That is, a node N_1 c-commands a node N_2 if the first branching node that dominates N_1 also dominates N_2 (Haegeman, 1994).



In the above tree, *A* c-commands all the nodes in the tree except *M*. *A* c-commands *C*, for instance, because the first branching node that directly dominates *A* (*M*) also dominates *C* whereas *C* does not c-command *A* because the first branching node dominating *C* (*B*) does not dominate *A.* In short, a node c-commands its sister node and all the children of the sister node that it dominates, according to standard definitions (Radford, 2004).

2.2.2.2 Properties of reflexives in English.

An English reflexive (e.g., herself) is a morphologically complex anaphoric pronoun composed of a pronominal and a root *-self* and should be coreferential with a previously mentioned nominal (its antecedent) within the same clause. More specifically, reflexives in English are bound by subject antecedents (subject orientation) and require their antecedent to fall within the same binding category (i.e., local clause). Moreover, they must be bound by a c-commanding antecedent. An item that has matching features with the reflexive cannot act as the antecedent unless these conditions are met. For instance, both noun phrases 'the woman' and 'the girl' in (8) are within the same clause of the reflexive and match it in features. However, only the noun phrase 'the mother' and not 'the schoolgirl' can be the reflexive's antecedent, because the former noun phrase c-commands the reflexive while the latter does not.

(8) **The woman**$_j$ behind the girl$_i$ hurt **herself**$_{*i/j}$.

In sum, to establish a correct referential link between an English reflexive and an antecedent, the reflexive must be c-commanded by a nominal referent from the same local clause.

2.2.2.3 Properties of reflexives in Arabic.

Reflexives in MSA are morphologically complex and they consist of a root *nafs* (soul or self) and a pronominal suffix that corresponds to person, gender and number (e.g. *nafs-hu*(3rdPr,SG.masc.)=himself, *nafs-ha*(3rdPr,SG.fem.)=herself, *anfs-hum*(3rdPr, PL.masc.)=themselves), which is relatively not different from English except for the different morphological arrangement of the reflexive word itself (AlBzour, 2015). In addition, they are similar in both semantic and syntactic functions as they behave in much the same way in relation

to antecedent orientation, locality domain property and c-command condition. That is, Arabic reflexives are locally bound and the antecedent should be the closest c-commanding nominal (subject) to the reflexive agreeing in person, gender and number, as illustrated in the following example (AlKafri, 2013).

(9) **alwaladu** alwaqıf khalf alrajuli dʒraħ **nafs-hu**

The boy-3M.SG. behind the man-3M.SG. hurt **himself**-3M.SG

"**The boy**$_j$ behind the man$_i$ hurt **himself**$_{*i/j}$"

In (9), both NPs (alwaladuu 'the boy') and (alrajuli 'the man') are within the reflexive's local domain and share the same features in person, number, and gender. However, 'alrajuli' cannot be coreferential with the reflexive as it does not c-command it, so binding in this case is not possible and the reflexive can only be bound by (alwaladu) as per Principle A. That is, c-command is the criterion used to distinguish eligible from ineligible and NPs.

Overall, subject-verb agreement and reflexive antecedent dependencies in Arabic and English share a number of key features. In subject-verb agreement, morphological features are taken into account in both languages and the verb should only be matched with its grammatical controller 'the subject'. Though the verb in MSA appears singular with dual and plural subjects disagreeing with them in number, this is only seen in VSO word order, whereas SVO word order, which parallels English sentence's word order, requires full agreement between the subject and verb in all features including number. Moreover, many local Arabic varieties, including Saudi Arabic that participants in the current studies speak, do not have this partial agreement and the verb agrees with the subject in number regardless of its number value and position. Likewise,

reflexives behave similarly in both English and Arabic. They are subject-oriented and require an antecedent that they agree with in terms of person, number and gender features. More importantly, they are constrained by Principle A in the sense that they must be bound by the closest c-commanding antecedent located within the same local binding domain.

Due to these similarities between the two languages in the relevant structures, L1 transfer that may occur when two languages are dissimilar and cause errors or processing difficulties is likely to be ruled out. If there are any processing differences between English and Arabic speakers, they would be better explained by other factors than L1 transfer.

2.3 Cue-based Memory Retrieval in L1 Sentence Processing

In recent years, there has been an increasing interest in examining working memory operations as an attempt to understand the mechanisms used by parsers in sentence comprehension. It is generally agreed that successful comprehension requires working memory to carry out necessary computational operations that involve encoding information, making connections and then retrieving and drawing conclusions. In this regard, a number of researchers in sentence comprehension have argued that processing non-adjacent dependencies relies on *a cue-based memory retrieval mechanism* (Lewis & Vasishth, 2005; Lewis et al., 2006; McElree, 2000; McElree et al., 2003). The main concept of cue-based parsing is that each incoming word during incremental parsing triggers retrieval to link that word with the preceding structure taking into consideration a group of features of the item known as *retrieval cues*. The following example of filler-gap dependency (adapted from Cunnings, 2017) illustrates how the relationship between the verb 'cleaned' and the dislocated filler 'the kitchen' as its direct object is established via cue-based parsing.

(10) It was **the kitchen** that the maid **cleaned**.

The result of processing the filler 'the kitchen' in (10) when it is first encountered is an encoded memory representation that will be stored in memory while the other words are being processed. At the time of reaching the verb 'cleaned', a retrieval of the filler as the direct object of the verb will be triggered via a retrieval operation.

As indicated earlier, cue-based parsing naturally occurs in stages of encoding, storing and retrieval. That is, linguistic units are firstly represented and stored in memory as chunks or feature

bundles with a set of feature values (Anderson et al., 2004). Those features include lexical content (e.g., morphological features, category, etc.) and structural relations (e.g., grammatical roles) working as symbols or pointers to other related chunks. This implies that parsing a sentence is basically a collection of chunks that are interlinked by feature specifications (Lewis & Vasishth, 2005).

After the encoding, the retrieval phase then is accomplished by accessing memory representations using retrieval cues as a subset of the features of the item to be retrieved (Lewis et al., 2006). For instance, to successfully comprehend sentences like (10), a set of syntactic and/or semantic retrieval cues like [+OBJECT], [+NP] and maybe [+CLEANABLE] can be derived from the verb and local syntactic context and used to guide the retrieval process to retrieve the target item in memory that can act as the object of the verb 'cleaned'. This retrieval process operates a *direct-access retrieval* mechanism, meaning that the retrieval cues are matched against all possible chunks in memory in parallel until the best match of cue features is located. It is also perceived as a content-addressable retrieval operation since cues enable direct access to relevant representations in memory based on their content[6] (Foraker & McElree, 2011; McElree, 2000; McElree et al., 2003; Van Dyke & McElree, 2011). However, a notable aspect of this memory retrieval process is its potentiality to generate similarity-based interference effects as a result of comparing retrieval cues with all items in memory, as will be further discussed in the next section.

2.3.1 Similarity-based interference.

Similarity-based interference is the hallmark of direct-access memory retrieval. A large body of research has shown that the presence of similar items in memory results in retrieval errors,

[6] Precise characterisation of the memory retrieval mechanism in sentence processing has previously been debated about whether the retrieval process is a direct-access operation or involves a serial search. Empirical evidence has shown that a direct-access mechanism is implemented in a number of retrieval tasks (e.g. Martin & McElree, 2008, 2009, 2011; Van Dyke & McElree, 2011) and the existing findings are best captured by the direct-access retrieval method (see Parker et al., 2017 for discussion). However, this issue is not within the scope of the current investigation.

decreasing the possibility of retrieving the target item (Lewis et al., 2006; Nicenboim & Vasishth, 2018; Parker et al., 2017; Vasishth et al., 2019). This phenomenon is known as similarity-based interference and has been investigated in many L1 processing studies (see Jäger et al., 2017 for a meta-analysis). From the cue-based retrieval perspective, two opposing interference effects may occur as a result of similarity-based interference: *inhibition* and *facilitation;* each is anticipated to arise in a different condition as illustrated below.

2.3.1.1 Inhibitory interference.

Inhibitory interference occurs in grammatical sentences when access to the target item is disrupted by other distractors that also match a subset of the retrieval cues. This results in slower reading times and lower accuracy during sentence comprehension (Parker et al., 2017). The overlap in feature content between distractors and a target can decrease the target's distinctiveness in memory, which may inhibit access to it. Increasing the number of distractors or the degree of their similarity, either semantically or syntactically, to the target would likely increase the difficulty. For instance, to understand who was doing the action (complaining) in the following example, the verb phrase needs to be attached to a noun phrase that is grammatically a local subject of the clause in which the verb is in and should also be animate. Accordingly, the retrieval cues [+SUBJECT] and [+ANIMATE] at the verb 'was complaining' could be used to retrieve the target noun phrase 'the resident'.

(11) a. The worker was surprised that **the resident** who was living near the dangerous neighbour **was complaining** about the investigation.

b. The worker was surprised that **the resident** who was living near the dangerous warehouse **was complaining** about the investigation.

In (11a), however, there are two noun phrases in the same clause 'resident', which is *the target*, having *the features* of $^{+\text{SUBJECT}}$ and $^{+\text{ANIMATE}}$, and 'neighbour', which is *the distractor*, with *the features* of $^{-\text{SUBJECT and }+\text{ANIMATE}}$, so both share the same $^{+\text{ANIMATE}}$ feature. Van Dyke (2007), as a result, observed that *inhibitory interference* appeared in sentences like (11a) compared to (11b) where the distractor 'warehouse' has $^{-\text{ANIMATE}}$ feature. That is, when the semantic features of the distractor 'neighbour' in (11a) fit the semantic retrieval cue from the verb, longer reading times were observed in condition (11a) relative to (11b). This suggests that the attention to retrieve the target noun phrase has been distracted due to the distractor that partially matched the retrieval cues, leading to decreased accuracy and longer retrieval time at the verb.

Several studies have found evidence for inhibitory interference of semantic and syntactic cues in grammatical non-adjacent dependencies (Van Dyke, 2007; Van Dyke, Johns, & Kukona, 2014; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006, 2011). Van Dyke and colleagues have supported the role of inhibitory effects in several studies. As shown in her study (Van Dyke, 2007), the semantic match of the animacy feature produced inhibitory interference, and the same result was found in Van Dyke & McElree (2011) during processing similar configurations. However, inhibitory effects have also been observed due to syntactic matching (Van Dyke & Lewis, 2003). As an example, (12b) caused more difficulty and required longer reading times than (12a) because the distractor noun phrase 'warehouse' in (12b) is syntactically the subject in the intervening relative clause. This matches the syntactic retrieval cue at the verb 'complained', which searches for an NP subject to complete the dependency. However, it is the object of the prepositional phrase in (12a).

(12) a. **The resident** who was living near the dangerous warehouse **complained** about the noise.

b. **The resident** who declared that the warehouse was dangerous **complained** about the noise.

Van Dyke and McElree (2006), in addition, found that processing filler-gap dependencies with clefted objects can be subject to inhibitory interference when sentences were preceded by items that are semantically associated with the verb of the sentence. Before reading sentences, participants were presented with a memory load of items and asked to remember them while reading as in (13).

(13) a. [table-sink-truck] It was **the boat** that the guy who lived by the sea **sailed** in two sunny days.

b. [table-sink-truck] It was **the boat** that the guy who lived by the sea **fixed** in two sunny days.

In contrast to the verb 'sailed' in (13a), the memory load list's items have a semantic association with the verb 'fixed' in (13b) since they all can be fixed. Consequently, participants' reading times were longer in (13b) than in (13a). This indicates that the presence of multiple similar NPs in memory has increased processing difficulty as the retrieval cues used for finding the target might have been disrupted with other items in memory impeding smooth elicitation of the target. This effect, however, was not noticed when they read conditions with no memory load

items, suggesting that retrieval interference could be behind the increased reading time rather than sentence complexity.

Furthermore, a recent meta-analysis by Jäger et al. (2017) included sentence comprehension experiments that investigated interference in un/grammatical dependencies, where grammaticality is based on the target's match or mismatch with the retrieval cues and interference occurrence was determined by the distractor match with any sort of the possible retrieval cues (animacy, number, etc.). The qualitative literature review showed that inhibitory interference was consistently observed, in line with the cue-based memory retrieval predictions, in grammatical non-agreement subject-verb dependencies as shown above in sentences similar to (11/12) (Van Dyke, 2007; Van Dyke et al., 2014; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006, 2011), but only occasionally in reflexive-/reciprocal-antecedent dependencies (Badecker & Straub, 2002, Experiment 3 & 4; Clackson & Heyer, 2014; Jäger, Engelmann, & Vasishth, 2015, Experiment 2; King, Andrews, & Wagers, 2012; Parker & Phillips, 2014; Patil, Vasishth, & Lewis, 2016) as most of the studies could not find conclusive results for such effects (Dillon et al., 2013; Sturt, 2003). Inhibitory interference was not also well-attested in S-V agreement studies. Nicenboim, Vasishth, Engelmann, & Suckow (2018), however, could find weak evidence of inhibitory effects as a result of number interference in German arguing that high-powered studies are vital to detect such effects.

The Bayesian quantitative analysis of Jäger et al. (2017) confirmed the inhibitory interference documented in the qualitative analysis for non-agreement subject-verb dependencies, as predicted by the cue-based retrieval account. However, there was no evidence of inhibitory interference in reciprocal-/reflexive-antecedent dependencies or S-V agreement dependencies, contradicting the cue-based retrieval account. The discrepant results were attributed to the different materials used in the studies and some methodological issues.

2.3.1.2 Facilitatory interference.

In addition to the prediction of increased processing difficulty when matching distractors are present along with the grammatical target in grammatical sentences, namely inhibitory effects, interference also increases the probability of incorrectly retrieving a partially matching distractor in ungrammatical sentences when the grammatical item is unavailable to reduce the ungrammaticality effect *(grammatical illusion)*. That is, the parser may perceive an ungrammatical sentence as grammatical due to a matching distractor's interference (Hammerly, Staub, & Dillon, 2019; Lewis & Phillips, 2014). This kind of interference is known as *facilitatory interference*[7] and is manifested as a speed-up in reading time (Parker et al., 2017). It occurs in ungrammatical conditions when the presence of a grammatically illicit but feature-matching distractor facilitates the cost of processing a grammatically inaccessible item (Engelmann, 2016).

Facilitatory interference has mostly been witnessed in literature in subject-verb agreement dependencies (Dillon et al., 2013; Lago, Shalom, Sigman, Lau, & Phillips, 2015; Parker & Phillips, 2017; Pearlmutter et al., 1999; Schlueter, Williams, & Lau, 2018; Tucker, Idrissi, & Almeida, 2015; Wagers, Lau, & Phillips, 2009) and negative polarity items licensing (Parker & Phillips, 2016; Vasishth, Brussow, Lewis, & Drenhaus, 2008). The evidence of facilitatory interference can be clearly illustrated by the following example.

(14) a. **The key** to the cabinet(s) **was** rusty for many years of disuse.

b. ***The key** to the cabinet(s) **were** rusty for many years of disuse.

[7] Facilitatory interference is sometimes called "attraction", as in the literature on S-V agreement attraction, or "intrusion". However, facilitatory interference is the general term used since they all have the same processing effects (Parker et al., 2017).

In the ungrammatical sentence (14b) in which the grammatical subject 'the key' mismatches the verb 'were', Wagers et al. (2009) noticed facilitated reading times after the verb and increased acceptability when the verb 'were' agrees in number with the distractor 'the cabinets' compared to the condition when the distractor 'the cabinet' is singular. The facilitation observed was attributed to a misretrieval of the plural distractor that matches the [+PLURAL] retrieval cue at the verb, as predicted by the cue-based memory retrieval model. That is, when the retrieval cues at the verb 'were' trigger a retrieval process to find a noun phrase in memory that has $^{+\text{SUBJECT}}$ and $^{+\text{PLURAL}}$ features, the distractor might be retrieved in some proportion of trials due to the partial match with the [+PLURAL] cue leading to the false conception that agreement is licit and resulting in facilitated processing relative to conditions where there is no plural distractor. No differences, however, were found in reading times between the grammatical conditions in (14a), though inhibitory interference is also expected in the grammatical condition with a matching distractor according to the cue-based model's predictions. Wagers et al. (2009) suggested that the distractor is less likely to be misretrieved in grammatical conditions because the complete match of the target is expected to outcompete the partial match of the distractor, and cue-based memory retrieval interference effects may accordingly occur only in ungrammatical conditions.

Similarly, grammatical illusions have been shown to arise during the processing of negative polarity items (NPIs) (Vasishth et al., 2008; Xiang, Dillon, & Phillips, 2009; Parker & Phillips, 2016). NPIs like 'any' and 'ever' should occur in a specific semantic context (e.g., negation) and within the licensed domain where it is c-commanded by a negative element as illustrated in (15a).

(15) a. **No** professor will **ever** say that.
b. *A professor will ever say that.
c. *The professor that no student likes will ever say that.

(15b) should be treated as unacceptable since the NPI is not in the scope of a negative element. (15c) should not also be accepted because the potential unacceptable NPI licensor is a subject of a relative clause that does not c-command 'ever'. However, the results frequently showed intrusive licensing effects in sentences like (15c) where a negative item precedes the NPI but does not c-command it. As a result, reading times in (15c) were faster than in (15b) due to the presence of a potential licensing NPI. It has, therefore, been argued that the erroneous retrieval of the grammatically unacceptable negation that partially matches the retrieval cues at the NPI, namely the [+NEGATION] cue, caused this grammatical illusion that facilitated its processing (Vasishth et al., 2008).

Jäger and colleagues' (2017) meta-analysis, both qualitative and quantitative, also showed strong evidence for facilitatory interference especially in S-V agreement dependencies regardless of the distractor's position (Dillon et al., 2013; Lago et al., 2015; Parker & Phillips, 2017; Tucker et al., 2015; Wagers et al., 2009). However, a considerable amount of variability was reported in the qualitative literature review regarding interference effects in reciprocal-/reflexive antecedent dependencies (Jäger et al., 2017). Several reflexive processing studies did not find evident support for facilitatory interference in ungrammatical conditions (Dillon et al., 2013; Sturt, 2003) or only marginally significant effects (Patil et al., 2016) or selectively in configurations where the target mismatches the anaphor in two features (e.g., gender and animacy) rather than only in one feature (Parker & Phillips, 2017). The quantitative analysis did not also show evidence for facilitatory interference in reciprocal-/reflexive antecedent dependencies. However, Jäger and colleagues' (2017) meta-analysis should not be taken as definitive due to its limitations. For example, the regions from which reading times were considered in the qualitative review differ from those used in the quantitative analysis. Moreover, the unavailability of raw data from some studies has forced the researchers to rely on the published numbers of those studies. The different sample sizes and

the low statistical power in experiments, along with the inconsistency of materials used across different studies, can also pose problems. All of these factors could lead to a discrepancy between the qualitative and qualitative analysis, making it difficult to draw strong conclusions.

Overall, cue-based parsing can result in two kinds of interference that affect reading times in different ways. The first one refers to inhibitory interference which may occur in grammatical (target-match) conditions as a result of an overlap in feature content between the target and the distractor, causing slower reading times and increased processing difficulty. The second refers to facilitatory interference that occasionally arises in ungrammatical (target-mismatch) conditions due to interference when a feature-matching but grammatically illicit item is misretrieved to facilitate processing. Despite the variability of findings, some empirical evidence could have been seen for both facilitatory and inhibitory interference in sentence processing in line with the cue-based memory retrieval account's predictions, though significant effects have not been reported for both in a single study (Jäger et al., 2017).

Having discussed different kinds of interference that may arise during sentence comprehension as a result of the cue-based memory retrieval process, I will move on to review a number of L1 reading-time studies that investigated the real-time processing of English S-V agreement and reflexive-antecedent dependencies from the interference-based perspective.

2.4 Interference in L1 Processing of Discontinuous Dependencies

As the focus of this thesis is on S-V agreement and reflexive resolution, below I discuss the L1 literature on these two dependencies in more detail.

2.4.1 Processing subject-verb agreement dependency.

Across a series of self-based reading experiments, Wagers et al. (2009) examined the effects of distractor number on native English speakers' reading times for grammatical and ungrammatical sentences in agreement attraction environments like (14), reproduced here as (16).

(16) The key to the cell(s) unsurprisingly (was/*were) rusty from many years of disuse.

In various experiments that utilized different types of verbs (copula vs. main verbs) and construction types with distractors placed in different positions, participants demonstrated sensitivity to grammaticality. Importantly, the presence of a distractor that matched the number of the verb resulted in reduced ungrammaticality effects and decreased reading times. Since facilitatory effects could also be noticed even if the distractor does not occur between the target subject and the verb as in 'the musicians who the reviewer praise...', this distinguishes it, on the other hand, from proximity concord in which a verb agrees with the closest NP rather than its subject NP (Quirk, Greenbaum, Leech, & Svartvik, 1985). There was, however, no evidence of a slowdown in reading times due to attraction effects from the distractor number in the grammatical conditions. Additionally, the attraction effects that were observed in ungrammatical sentences were limited to instances where a plural distractor was present. This finding replicates the same agreement attraction pattern, the mismatch asymmetry, observed solely for singular subject-plural distractor conditions but not plural subject-single distractor conditions in previous literature

(Bock & Miller, 1991; Pearlmutter et al., 1999). This was attributed to the explicitly marked number feature of the plural compared to singulars, which in turn leads to more attraction.

A final offline acceptability judgement experiment further assured the observed attraction pattern. A plural distractor led ungrammatical sentences to be accepted but did not cause grammatical sentences to be rejected. Wagers and colleagues consequently suggested that the asymmetrical pattern of attraction effects in grammatical and ungrammatical sentences can result from a cue-based retrieval mechanism. That is, when the retrieval cues at the verb are used to retrieve the subject from memory, a fully matching noun is more likely to be retrieved than a partially matching noun. This is clearly the case in grammatical conditions where no attraction effects were observed. However, when neither of the nouns matches all the cues, the number-matching illicit noun can sometimes be retrieved, as is the case for ungrammatical conditions where attraction usually arises. Another explanation offered for the asymmetric attraction effects is that the moment the mismatch between the subject and the verb is realised during parsing, this error signal initiates a reanalysis process in which retrieval cues are then used to find a matching antecedent to eliminate the mismatch effect. Both options yet rely on a cue-based retrieval mechanism. Overall, different experiments suggest that processing S-V agreement can be susceptible to facilitatory interference but not inhibitory interference.

Parker and An (2018) examined S-V agreement dependencies in three self-paced reading experiments mainly to test Van Dyke and McElree's (2011) proposal that interference is mediated by the distractor's argument status. Specifically, they hypothesized that core argument distractors (e.g., direct objects and subjects), which are necessary to establish the meaning of the sentence, are distinctively encoded in memory relative to oblique arguments, as prepositional phrases (PP), and hence are more resistant to interference. On the basis of this account, Parker and An (2018) initially compared direct object and PP distractors as in (17).

(17) **a. PP distractor**

The waitress who sat near the girl(s) unsurprisingly (was/*were) unhappy…

b. Direct object distractor

The waitress who sat the girl(s) unsurprisingly (was/*were) unhappy…

The results revealed a significant main effect of grammaticality in both distractor conditions, with longer reading times observed for ungrammatical sentences. However, while a main effect of distractor was observed, along with a significant interaction between grammaticality and distractor in PP distractor conditions, no corresponding evidence of interference was observed in direct object distractor conditions. The findings supported Van Dyke and McElree's (2011) proposal regarding the difference between the direct object and PP distractors in their susceptibility to interference. It was, therefore, argued that direct objects are less expected to interfere at retrieval because they are easier to reject when they mismatch the syntactic retrieval cue since they are more distinct in memory than PPs.

In their second experiment, Parker and An (2018) examined Van Dyke and McElree's (2011) account of encoding by comparing direct object and subject distractors as in (18).

(18) **a. Direct object distractor**

The celebrity who insulted the journalist(s) certainly (was/*were) upset …

b. Subject distractor

The celebrity who the journalist(s) insulted certainly (was/*were) upset …

Interference was predicted to most likely occur when the distractor is in a subject position because of its match with the syntactic retrieval cue. However, the results revealed that both conditions showed significant grammaticality effects, but no distractor or interference effects were observed in any region for both conditions. Experiment 2 replicated Experiment 1 results in showing that direct objects resist interference.

Parker and An (2018) concluded that PP distractors beat object and subject distractors with regards to triggering interference, in consistence with Van Dyke and McElree's (2011) proposal that was primarily based on their study of subject-verb thematic binding and extended by Parker and An (2018) in showing that similar patterns could be observed in subject-verb agreement. Taken together, several studies in the sentence comprehension literature have shown evidence of interference effects yielding facilitated processing in ungrammatical subject-verb agreement sentences due to the presence of an inappropriate but number-matching NP (Dillon et al., 2013; Lago et al., 2015; Parker & Phillips, 2017; Wagers et al., 2009), though the syntactic position of a distractor may influence facilitatory interference (Parker & An, 2018). No reliable corresponding interference effects yielding processing costs in grammatical sentences have been observed, however.

It is noteworthy that facilitatory interference should not be seen as a misrepresentation or distortion of the subject's number properties, as has previously been argued, especially for attraction effects in production (Bergen & Gibson, 2012; Eberhard, Cutting, & Bock, 2005). Models of representation distortion accounts typically argue that attraction arises from a faulty or

unclear representation of the subject, and they can be divided into two types (Bergen & Gibson, 2012; Eberhard et al., 2005; Franck, Vigliocco, & Nicol, 2002; Hammerly et al., 2019). According to the percolation model (Bock & Eberhard, 1993; Eberhard, 1997; Franck et al., 2002), attraction occurs due to a faulty upward movement of the distractor's marked plural feature to the subject during processing the subject NP, where it can be copied onto the verb. Therefore, similar attraction effects are also predicted to occur in grammatical conditions like 'The key to the cabinets was…' and not only in ungrammatical conditions like 'The key to the cabinets were…'. In the Marking and Morphing model (Eberhard et al., 2005), attraction occurs due to equivocal or ambiguous representation of the complex subject's number in conditions with distractors that mismatch the agreement controller's number 'e.g., the key to the cabinets…' and predicts that this equivocal number marking would affect both grammatical and ungrammatical conditions, similarly to the percolation model's prediction. However, while facilitatory interference or the so-called attraction has usually been observed in ungrammatical conditions, parallel effects have rarely been reported in grammatical conditions. This does not indicate that there is no existing data on attraction symmetry in grammatical and ungrammatical sentences (e.g. Pearlmutter et al., 1999, Experiment 1 and 2), but they are much less and inconsistent. Wagers et al. (2009), however, noted that the processing difficulty observed in most of these cases at the verb region in grammatical-plural distractor conditions represents plural NP spillover effects or an additional processing load of the plural distractor that immediately precedes the verb since plurals are morphologically and arguably conceptually more complex than singulars. That is, when the distractor and verb are adjacent, this design confound may cause the observed effects. Accordingly, this grammatical asymmetry observed in agreement attraction has been taken to argue in favour of cue-based parsing as evidenced by the facilitatory interference detected during

agreement processing (Wagers et al., 2009)[8]. That is, the facilitation effect observed was considered to be the behavioural signature of the cue-based retrieval mechanism.

2.4.2 Processing reflexive-antecedent dependency.

A number of studies investigated the interference effects of grammatically illicit NPs during antecedent retrieval in reflexive-antecedent dependencies as a useful tool for understanding processing mechanisms. Previous research was particularly focused on investigating the types of information utilised during online reflexive interpretation. This includes whether structural constraints are only involved in coreference processing or if other non-structural information can also interact during reflexive interpretation. To explore this, an interference paradigm has typically been utilised, which systematically varies non-structural information.

Badecker and Straub's (2002) study is one of the earliest studies that investigated whether the content of the structurally illicit NPs may influence the processing of reflexives or that the Binding Principle would operate solely and quickly eliminating any other, even matching, grammatically illicit nouns from being considered as potential antecedents for the reflexive. To this end, the gender of the structurally illicit subject noun was manipulated as in (20).

(20) a. Jane thought that Bell owed himself another opportunity to solve the problem.

b. John thought that Bell owed himself another opportunity to solve the problem.

[8] Based on a quantitative computational modelling comparison of multiple models to explain number agreement patterns observed in published studies, Yadav, Smith, Reich, and Vasishth (2023) recently proposed a novel model of agreement attraction that assumes a hybrid mechanism incorporating both representation distortion and retrieval processes such that the cue-based memory retrieval process operates on the possibly distorted representations of nouns in memory. However, we do not intend to tease these accounts apart, but rather examine the degree to which memory retrieval interference that may result from cue-based retrieval, as a general theory of dependency formation, impacts the processing of L1 and L2 speakers.

The results revealed that native English speakers' reading times were affected by the gender compatibility of the structurally illicit NP in the form of longer reading times in (20b) that contained a gender-matching distractor compared to (20a) that did not. This evidence represents inhibitory interference as described in the cue-based retrieval model. It suggests that reflexive resolution is not only mediated by structural constraints, but other sorts of information can also interact during online reflexive interpretation.

Similarly, Sturt's (2003) prominent eye-tracking study on reflexive processing examined native English speakers' sensitivity to grammatically illicit but feature-matching distractors during reflexive antecedent retrieval. As shown below, the gender match between the target local gender-stereotyped NP 'e.g., the surgeon' and the non-local proper name distractor 'Jennifer or Jonathon' was varied.

(21) **a. Target-match/Distractor-mis/match**

Jennifer/Jonathan was pretty worried at the City Hospital. She/He remembered that the surgeon had pricked himself with a used syringe needle. There should be an investigation soon.

b. Target-mismatch/Distractor-mis/match

Jonathan/Jennifer was pretty worried at the City Hospital. He/She remembered that the surgeon had pricked herself with a used syringe needle. There should be an investigation soon.

The results of early measures showed that reading times were faster at the reflexive region

when the target antecedent's gender 'the surgeon' matched the reflexive (e.g., surgeon …himself) than when it did not (e.g., surgeon…herself) indicating significant grammaticality effects. Moreover, reading times did not significantly differ due to the distractor gender match with the reflexive. The results of later measures, however, revealed the distractor's influence in conditions where the target matched the reflexive. That is, reading times increased when the distractor gender did not match the reflexive's compared to when it did. In a follow-up task that aimed to check participants' interpretation of the sentences by asking directly about the reflexive's antecedent, the proportion of ungrammatical responses, where participants chose the distractor as the antecedent, was calculated. The results showed a main effect of the distractor as the ungrammatical responses were more prevalent in the distractor match conditions than in the distractor mismatch ones. Further confirmation of the distractor's influence was the fact that the target-mismatch/distractor-mismatch condition had the lowest proportion of ungrammatical responses.

To confirm that the early effects observed for the target antecedent were not caused by its linear position to the reflexive, Sturt conducted a second experiment in which the linear position of both antecedents to the reflexive was altered, as shown in (22).

(22) Jennifer/John was pretty worried at the City Hospital. The surgeon who treated Jennifer had pricked himself/ herself with a used syringe needle. There should be an investigation soon.

The results showed only a significant main effect for the target antecedent with no significant effects for the distractor whether in late or early measures, in contrast to Experiment

1. Sturt concluded that participants' processing had an overall grammatical preference, and the distractor effects observed occurred considerably later than the target's effects. Insufficient justification, however, was given to account for the responses of the comprehension questions.

In contrast to Sturt's (2003) study, Patil et al. (2016) found evidence of the use of non-structural cues, particularly gender, during the initial retrieval of the reflexive antecedent. They adapted Sturt's (2003) Experiment 2 by placing the distractor in a subject position rather than in an object position within an embedded clause, as in (23), arguing that this would result in a stronger interference effect.

(23) **a. Target-match/Distractor-mis/match**

The tough soldier that Katie/Fred treated in the military hospital introduced himself to all the nurses.

b. Target-mismatch/Distractor-mis/match

The tough soldier that Fred/Katie treated in the military hospital introduced herself to all the nurses.

The results showed that reading times increased when the distractor matched the reflexive's gender in grammatical conditions, indicating early interference effects. However, only late effects of the target antecedent were observed when it did not match the reflexive's gender. This goes against Sturt's (2003) conclusion that interference effects cannot arise at the early stages of processing. Besides this inhibitory interference reported in grammatical conditions, a marginally significant facilitatory interference was also found in ungrammatical conditions with

matching distractors supporting cue-based retrieval's predictions.

In a series of eye-tracking experiments, Parker and Phillips (2017), compared contexts where the target and reflexive were mismatched in one feature (e.g., gender, etc.) and in two features (e.g., gender and animacy). However, no evidence of facilitatory interference from a matching distractor could be found unless the target and reflexive were mismatched in two features. It was consequently suggested that reflexives are prone to interference, but only selectively. That is, interference can arise when the reflexive and target mismatch in multiple features but not when they mismatch with only one feature.

Taken together, though the findings reviewed above are mixed, some studies indicate that native English speakers may integrate syntactic knowledge with other non-structural information during referential processing which can sometimes cause processing disruptions, as demonstrated by interference effects.

2.4.3 Comparing agreement and reflexives.

One theoretical issue discussed in some L1 processing studies is whether different grammatical dependencies show the same pattern of sensitivity to different linguistic features when retrieving the target item for dependency resolution (Dillon et al., 2013; Jäger et al., 2020; Parker & Phillips, 2017). In a study comparing S-V agreement and reflexive-antecedent dependencies processing in English, Dillon et al. (2013) manipulated the distractor number feature in reflexive conditions like (24) and in S-V agreement conditions like (25) to investigate any potential differences in interference profiles for these dependencies.

(24) Agreement

a. The new executive who oversaw the middle manager(s) apparently was dishonest about the company's profits.

b. *The new executive who oversaw the middle manager(s) apparently were dishonest about the company's profits.

(25) **Reflexives**

a. The new executive who oversaw the middle manager(s) apparently doubted himself on most major decisions.

b. *The new executive who oversaw the middle manager(s) apparently doubted themselves on most major decisions.

The results revealed a significant main effect of grammaticality in both dependencies, with slower reading times for ungrammatical than grammatical sentences. Reading times also showed only evidence of facilitatory interference effects in agreement ungrammatical conditions. However, no significant interference effects were found in reflexive dependencies whether in grammatical or ungrammatical conditions. Dillon and colleagues, accordingly, argued that surface agreement features are not similarly deployed during retrieval as the lack of interference effects in reflexives indicates that reflexive antecedent retrieval involves only structural cues, unlike S-V agreement where overt agreement features are also engaged during retrieval.

A recent replication of Dillon et al. (2013) study with a larger sample by Jäger and colleagues (2020), however, found similar facilitatory interference profiles between the two dependency types based on their analysis of total fixation time, the same measure Dillon and

colleagues drew their conclusion from. However, first-pass regression analysis showed clear facilitatory interference in S-V agreement but not in reflexive-antecedent dependency. By contrast, some indication of inhibitory interference was observed in reflexive but not S-V agreement conditions. Given that Dillon and colleagues did not find any interference in both dependency types in first-pass regressions and that very few studies have reported interference effects from first-pass regressions (see Jäger et al., 2017 for a meta-analysis), Jäger et al. were cautious to draw any strong conclusions about the effects they found at this measure. Finally, Jäger et al. argued that the interference profiles of both dependencies were generally similar, suggesting thus that the retrieval mechanisms employed by both dependencies are not distinct, as opposed to what Dillon and colleagues have suggested.

To sum up, it can be observed that facilitatory interference is typically observed in the L1 processing of S-V agreement, but is not consistently found during processing reflexive constructions. Inhibitory interference, on the other hand, has been reported in only a limited number of reflexive studies and has not been observed in S-V agreement studies. Additionally, no study has reported significant effects of both kinds of interference simultaneously (Jäger et al., 2017; cf. Jäger et al., 2020). Considering that interference effects in L2 processing, on the other hand, remain unclear, it would thus be interesting to see whether L2 processing of S-V agreement and reflexives also experiences similar effects or exhibits a greater susceptibility to interference (Cunnings, 2017).

2.5 Theoretical Approaches to L2 Sentence Processing

In real-time sentence comprehension in L2 readers, utilizing bottom-up and syntactic information along with lexical-semantic and discourse-pragmatic information while incrementally processing the language input is necessary to develop native-like behaviour. Whether L2 speakers can apply the same processing heuristics as L1 speakers has been debated, leading to different explanations for potential L1/L2 processing differences. The source of these differences has been located in either L2 speakers' reduced sensitivity to grammatical information (Clahsen & Felser, 2006a, 2018), individual differences and capacity-related aspects of language processing (Hopp, 2014, 2015, 2018; McDonald, 2006), or vulnerability to memory retrieval interference (Cunnings, 2017). The current research aims to investigate the L2 processing of Arabic L2 speakers of English in comparison to native English speakers in light of the main theoretical accounts discussed in the following sections.

2.5.1 The Shallow Structure Hypothesis.

As far as L2 processing is concerned, some studies in the existing literature focus particularly on the nature of the mechanism implemented by L2 speakers in L2 sentence comprehension compared to L1 speakers. Clahsen and Felser's (2006a, 2018) prominent study provided a landmark comprehensive account of L1/L2 parsing differences. According to their Shallow Structure Hypothesis (SSH), L2 speakers primarily perform meaning-based computation and are less dependent on grammatical information during L2 processing, compared to native speakers who rely more on abstract hierarchical syntactic structures. In other words, it has been suggested that the parsing problems faced by L2 speakers, even among highly proficient L2 speakers, especially during the processing of grammatically mediated discontinuous dependencies are because "L2 grammar does not provide the type of syntactic information required to process non-local grammatical phenomena in native-like ways" (Clahsen & Felser, 2006b, p. 565). It has also been argued that L2 speakers may engage in shallow processing, even

when the L2 grammar does not diverge from that of their L1, due to a lack of appropriate parsing heuristics necessary for processing the L2 language. This hypothesis has been supported by several studies that showed the influence of lexical-semantic properties on L2 speakers and their underuse of syntactic information during sentence processing (e.g., Felser & Roberts, 2007; Felser, Roberts, Marinis, & Gross, 2003; Marinis, Roberts, Felser, & Clahsen, 2005). For instance, Marinis et al. (2005) examined the processing of long-distance *wh*-dependencies in sentences like (26).

(26) a. The nurse [$_{CP}$ who$_i$ the doctor argued [$_{CP}$ e$_i$' that the rude patient had angered e$_i$]] is refusing to work late.

b. The nurse [$_{CP}$ who$_i$ the doctor's argument about the rude patient had angered e$_i$] is refusing to work late.

According to generative grammar (Chomsky, 1981, 1995), a dislocated constituent in dependencies that crosses more than a clause boundary should be mediated by empty syntactic categories at clause boundaries before linking it to its subcategorizer. That is, a long-distance dependency between a filler and its gap should not be formed in a single step, but rather in a series of small steps as per the Subjacency constraint. Both (26a/b) contain a long-distance dependency with the same length between the dislocated *wh*-constituent (or the filler) and its subcategorizer (or its ultimate gap) 'angered', but (26a) involves crossing a clause boundary unlike (26b). Hence, postulating an intermediate gap site (indexed by e$_i$') at the clause boundary during processing (26a) before integrating the *wh*-phrase with its actual gap position (e$_i$) is a syntactic requirement. If an intermediate gap is postulated and the filler is thus reactivated, reading times at the ultimate

gap position ‘angered’ in (26a) should be shorter than (26b) given the reduced distance between the filler and gap. Marinis et al. (2005), however, found that only L1 speakers showed effects of the subjacency constraint and the use of intermediate gap in their reading times, while L2 speakers integrated the filler directly with its lexical subcategorizer upon encountering it. This was interpreted to suggest that L2 processing of long-distance *wh*-dependencies lacks the use of abstract syntactic constraints, unlike L1 processing.

Most evidence of shallow processing was initially reported from processing relative clause ambiguities and filler-gap dependencies. Agreement processing, however, was also included in the scope as discussed in a commentary version of this hypothesis. A native-like sensitivity to agreement violations was argued to be possible for L2 speakers only “between closely adjacent constituents”, as in processing noun-adjective gender agreement within a noun phrase, but it is less likely to be achieved during processing S-V agreements with configurational structures (Clahsen & Felse, 2006, pp. 111, 115). In further elaboration, L2 speakers were also argued to have difficulty checking formal uninterpretable features, such as c-command successfully as in Chomsky’s (1995) view to process and relate, for example, an anaphor to its antecedent since this depends on hierarchical relationships between sentence constituents (Clahsen & Felse, 2006; Felser, 2019).

Other studies in the literature, however, showed that L2 speakers can involve syntactic computation during their processing of discontinuous dependencies and they behave in much the same way as native speakers with regards to applying syntactic constraints as opposed to what the SSH predicted (e.g., Aldwayan, Fiorentino, & Gabriele, 2010; Hopp, 2006; Omaki & Schulz, 2011; Pliatsikas & Marinis, 2013). Pliatsikas and Marinis’s study (2013), for example, contradicted previous studies (e.g., Marinis et al., 2005) which indicated that L2 speakers do not use intermediate gaps during long-distance dependencies processing. They, conversely, showed

that L2 speakers with sufficient naturalistic exposure to the L2 can establish native-like syntactic processing of long-distance *wh*-dependencies.

Thus far, the available evidence is inconsistent with some studies reporting L2 speakers' underuse of syntactic constraints and others finding them sensitive to those constraints. Earlier studies, however, generally built their arguments based on the type of information that L2 speakers might or might not recruit during processing to analyse L1/L2 parsing differences without much focus on individual differences that may have created different processing patterns across L2 processing studies. As such, L1/L2 processing differences have thus started to not be conceived by others (Hopp, 2010, 2014; McDonald, 2006) as an indication of qualitative differences in the parsing mechanism between native and non-native speakers, but rather as quantitative individual differences in processing resources or factors, such as L2 proficiency, lexical automaticity or working memory capacity, that might eliminate L1/L2 processing differences if they are taken into consideration. These will be discussed in the subsequent sections.

2.5.2 Proficiency in L2 sentence processing.

In research on L2 processing, proficiency has typically served as an inclusion criterion or included as an independent predictor, highlighting the pivotal role that general L2 proficiency plays in enhancing a native-like processing profile. Depending on the extent to which an individual knows linguistic representations and the ease with which they can process linguistic knowledge, syntactic processing might be affected.

Putting previous research into perspective, empirical evidence has shown that morphosyntax is increasingly used in incremental processing as proficiency advances. For instance, L2 speakers with low proficiency were found to be less sensitive to agreement violations (e.g., RT data: Keating, 2009; Lim & Christianson, 2014; Sagarra & Herschensohn, 2010, 2013; event-related potential (ERP) data: Ojima, Nakata, & Kakigi, 2005; Rossi, Gugler, Friederici, &

Hahne, 2006), suggesting that this sensitivity is partly determined by their level of proficiency in the second language. The study by Rossi and colleagues (2006), for example, found that native speakers and high proficiency L2 speakers in both L2 groups (L1 Italian speakers of German and L1 German speakers of Italian) exhibited the same ERP components; namely LAN and P600 effects, in response to S-V person agreement morphosyntactic violations, whereas low proficiency L2 speakers only showed a delayed and reduced P600 effect. Ojima et al. (2005) likewise noticed that semantic violations 'The house has ten rooms in total. vs. *The house has ten cities in total.' elicited the same brain responses as native English speakers in both higher proficiency and lower proficiency Japanese L2 speakers of English while syntactic agreement violations 'Turtles move slowly vs. *Turtles moves slowly' did not provoke similar effects in lower proficiency L2 speakers, indicating that the development of syntactic processing may rely more on proficiency relative to semantic processing.

Proficiency, however, seems to be less influential in real-time structural ambiguities processing relative to other types of individual differences or linguistic factors. Hopp (2006), though, found that the reanalysis processing patterns of L1 English and L1 Dutch speakers of German vary as a function of proficiency since only higher proficiency L2 speakers employ syntactic features and principles for reanalysis to a native-like degree. In a similar study, Hopp (2010) also noticed that proficiency affects the use of syntactic features during processing by different groups of L2 German speakers. However, no measurable effect of proficiency was observed in the online resolution of other structural ambiguities studies (e.g., Cheng, Rothman, & Cunnings, 2021; Cotter & Ferreira, 2022; Dussias & Sagarra, 2007; Hopp, 2014, 2015).

L2 speakers' proficiency level has rather been reported frequently as a criterion for inclusion in most early filler-gap dependencies studies (e.g., Felser & Roberts, 2007; Pliatsikas & Marinis, 2013) and the few studies on reflexive-antecedent processing (e.g., Felser et al., 2009), among other L2 processing studies. Testing a wider range of L2 proficiency, however, and

including it as a predictor would more likely show how the linguistic phenomenon under investigation interacts with proficiency and developmental progression.

Overall, several studies have shown evidence of proficiency's contribution to L2 syntactic processing, suggesting that some discrepancies between L1 and L2 processing can be accounted for by individual differences in L2 proficiency. It is thus worth testing L2 speakers' proficiency to see whether target-like sentence processing can ultimately be achieved as proficiency increases.

2.5.3 The Lexical Bottleneck Hypothesis.

Hopp's (2018) Lexical Bottleneck Hypothesis (LBH) provides a useful account of the role of lexical processing in sentence comprehension and shows how bilingual lexical processing characteristics may engender aspects of non-target language processing. It argues that L2 parsers do not qualitatively differ from L1 parsers, and the difficulties they face can be attributed to factors other than syntactic processing per se. Specifically, lexical processing, which is partly seen as a preceding stage that feeds into syntactic processing, can be affected by the bilingual mental lexicon which may, in turn, modulate L2 processing.

Since bilinguals use their languages less frequently than monolingual speakers, words are less activated, and lexical retrieval as a result tends to be slower. Evidence of slower lexical retrieval has been taken from L2 speakers' decision errors along with their slower responses in lexical decision tasks and their retrieval errors in picture naming (e.g., Gollan et al., 2011; Lemhöfer et al., 2008). Slower lexical retrieval among L2 speakers is also correlated with frequency effects, such that words with low frequency are processed more slowly by L2 speakers relative to native speakers. This indicates that associations between lexical forms and conceptual representations are reinforced through usage. Therefore, weaker lexical links, which are observed more with low-frequency items, can impact automatization in lexical processing resulting in

slower lexical processing which may in turn lead to different syntactic processing patterns compared to native processing.

Hopp (2018) thus highlighted the importance of taking individual differences in lexical processing into account for its potential contribution to syntactic processing. This has been observed for attachment preferences of relative clause ambiguities like 'The student had liked the secretary of the professor who had almost killed himself in the office'. Hopp (2014) found that L2 speakers with higher lexical automaticity demonstrated a significant native-like preference for low-attachment of the relative clause, whereas slower lexical decoders showed a null preference (Cheng et al., 2021). Similarly, Hopp (2015) noticed that L2 speakers with higher lexical automaticity showed early signs of using plausibility and verb transitivity information to reanalyse garden-path sentences than L2 speakers with less automatized lexical processing.

Not all phenomena, however, are expected to be affected to the same extent by lexical processing and lexical retrieval speed, and certainly, not all aspects of non-target syntactic processing can be accounted for by lexical retrieval difficulties. Nevertheless, it is useful to incorporate the LBH into existing L2 processing models. Hopp (2018), for instance, suggested that L2 speakers' underuse of syntactic information relative to L1 speakers, as argued by the Shallow Structure Hypothesis (2006a, 2018), can be a result of the weight they assign more to lexical information compared to syntactic information. That is, the taxing demands of lexical processing may contribute to the absence or delay of syntactic constraints' real-time application in L2 processing. Therefore, comparing L1 and L2 speakers in terms of their lexical processing would reveal whether they are influenced by lexical representations to the same extent and how similar they are if lexical factors are taken into account.

The LBH has been seen as a reflection of capacity-based models (e.g., McDonald, 2006), such that lexical processing may diminish resources used for sentence processing which in turn

may result in non-target-like syntactic processing[9]. The following section will also discuss working memory capacity effects, as another capacity-related aspect of language processing.

2.5.4 Working memory in L2 sentence processing.

2.5.4.1 Working memory capacity-based model.

The fact that sentence comprehension is essentially modulated by a cognitive system, namely working memory resources (WM), was taken as one of the potential explanations for the observed differences found in L2 processing studies. In practice, the effects of working memory capacity have been considered as a contributing factor in L2 processing, wherein limited cognitive resources may impede achieving native-like processing (Farmer, Misyak, & Christiansen, 2012). According to the capacity-based account[10] (Just & Carpenter, 1992; McDonald, 2006), the role that WM plays in sentence comprehension entails maintaining the representational elements constructed earlier in a sentence active in memory and executing a number of processes concurrently to integrate displaced constituents in a sentence during incremental parsing. The capacity-based model is thus mainly concerned with the *amount* of information that can be actively held or used at one time, or the *number* of processes that can be performed at once (Daneman & Carpenter, 1980; Just & Carpenter, 1992).

Previous research findings into the role of WM in L2 syntactic processing have also been inconsistent. Some studies, for instance, did not find reliable evidence of WM effects on L2 speakers' processing of filler-gap dependencies and ambiguous relative clauses (Felser & Roberts, 2007; Juffs, 2004; Nakano & Wang, 2011; Omaki, 2005). Others, however, found its

[9] Cunnings (2022), however, suggested that viewing the LBH from the capacity-based perspective needs further specification as the idea of how specific resources are assigned for lexical processing and others for other aspects of sentence processing (e.g., syntactic processing) is unclear. On this basis, it is also not clear how a parser's level of lexical access automaticity may curtail resources or make them available for other aspects of sentence parsing. The LBH can instead be viewed in terms of the type and quality of lexical representations that presumably provoke differences in the retrieval mechanism and eventually processing.

[10] Working memory has been conceptualized theoretically in different ways (Baddeley, 2007; Cowan, 1988; Engle, 2002), but the focus here will be on the capacity-based account as it is the prevalent model considered in sentence processing literature.

effects in some tasks but not in others (Hopp, 2014; Zhou, Rossi, & Chen, 2017). The role of memory span was also investigated in agreement processing. In their study of noun-adjective gender and number agreement processing by beginners and intermediate L1 English speakers of Spanish, Sagarra and Herschensohn (2010) found that WM influenced only intermediate L2 speakers' sensitivity to gender but not number agreement violations. Foote (2011) also examined noun-adjective gender and subject-verb number agreement in Spanish with manipulation of the distance between the noun/adjective and subject/verb to create conditions without/with an intervening material like, for example, (27a) and (27b), respectively for noun-adjective agreement.

(27) a. Dicen que el libro blanco/*blanca esta en esa mesa.

They say that the book-M.SG white-M.SG / *white-F.SG is on that table.

b. El pollo del taco está rico/*rica pero picante.

The chicken-M.SG of the taco is tasty-M.SG/ *tasty- F.SG but spicy.

L2 speakers of Spanish showed sensitivity to agreement errors, but that sensitivity was greater in the conditions without intervening material as in (27a) relative to the longer conditions as in (27b). Though memory span did not correlate with this sensitivity difference between the conditions as a result of distance, working memory limitations have still been suggested as a possible explanation that needs further research. Keating (2010) also found that noun-adjective agreement processing by advanced L2 speakers of Spanish was influenced by the distance between the noun and its adjective. L2 speakers' sensitivity to gender errors was also modulated by individual differences in WM.

Overall, the findings are rather controversial, and there is no consensus on the magnitude of working memory capacity's influence on L2 processing. Moreover, it should be noted that such a line of research has solely looked at working memory in terms of capacity limitations, without carefully considering the memory operations that underlie sentence processing. Despite that the present study is mainly formulated to examine the similarity-based interference account that may result from the cue-based memory retrieval mechanism as will be discussed below, we can still look into the capacity-based effects through the linguistic materials, though not in terms of individual differences in WM across participants.

2.5.4.2 Working memory capacity-based account vs. interference-based account.

Cunnings (2017, 2022) has recently argued that understanding how native and non-native parsers encode and retrieve information from memory can provide a nuanced explanation for L1/L2 processing differences rather than differences in working memory capacity. That is, interference during memory retrieval can better account for the problems attested in L2 processing. This view is based on the theoretical framework -cue-based parsing- which interprets sentence comprehension as a series of skilled memory retrievals (Lewis & Vasishth, 2005). As an illustration, in the following examples, Cunnings (2022) demonstrated how it is difficult to dissociate between capacity-based and interference-based accounts in sentence processing.

(28) a. The cat was scared.

b. The cat the dog the boy saw chased was scared.

(29) a. It was the book that the boy with the coffee very happily read before lunch.

b. It was the book that the boy with the magazine very happily read before lunch.

According to the capacity-based account, processing (28b) is predicted to be more difficult compared to (28a) given that 'the cat' and its predicate 'was scared' in (28b) are separated by

other constituents, unlike in (28a), and this may tax memory resources or exceed its capacity. The same kind of difficulty is also predicted by the interference-based account during processing (28b) relative to (28a), but the reason is attributed to potential interference from other constituents given that 'the dog' and 'the boy' could interfere as potential subjects of 'was scared'. Both accounts thus have the same prediction but offer different explanations, so dissociating these accounts empirically is difficult.

Cunnings (2022), however, pointed out that it is less obvious how differences between processing (29a) and (29b) could be explained by the capacity-based account given that they both have the same number of words but differ in regard to the meaning or type of words involved. For the interference-based account, processing (29b) would be more difficult. That is, when the noun phrase 'the book' would be retrieved as an object of the verb 'read', the noun phrase 'the magazine' in (29b) is more likely to interfere and could occasionally be retrieved as the object compared to 'the coffee' in (29a). This is because 'the magazine' is semantically more plausible as it refers to something [+READABLE], unlike 'the coffee'. In sum, while the amount of information and the ability to maintain representations active in memory is what may affect processing in sentence comprehension according to the capacity-based account, it is the type of content in memory and the ability to retrieve target information correctly that may influence comprehension according to the interference-based account (Parker et al., 2017).

According to Cunnings (2022), interference, therefore, cannot be ruled out in the studies of Keating (2010) and Foote (2011). For instance, 'taco' in (27b) might interfere when its gender matches the adjective 'rico'. Thus, further investigation to dissociate between the capacity-based and interference-based views is needed by systematically manipulating the intervening item (e.g., 'taco') to detect whether readers are susceptible to interference effects. That is, if readers in our studies, for example, are found to be influenced by the distractor only in conditions where interference is expected to arise, as per the cue-based parsing's predictions, we can argue in favour

of the interference-based account. If they, however, employ a recency match and relate the verb or reflexive to the closest noun phrase that has been stored more recently and hence is likely to be more active for retrieval, we cannot rule out the possibility that readers might be affected by WM capacity limitations.

Having discussed different conceptualisations of WM in sentence processing and other L2 theoretical accounts, the following sections will review sentence comprehension studies that have investigated L2 speakers' real-time processing of S-V agreement and reflexive-antecedent dependencies.

2.6 L2 Processing of Discontinuous Dependencies

2.6.1 L2 processing of subject-verb agreement dependency.

Previous studies on S-V agreement in L2 processing have reported mixed findings, with very few investigations addressing interference effects. According to Jiang's self-paced reading study (2004), Chinese L2 speakers of English were not sensitive to number disagreement between non-adjacent subjects and verbs compared to L1 speakers. It should be noted though that conclusions were drawn from analyses conducted for each group separately rather than including both groups into one analysis for direct comparison. Participants were also tested on grammatical conditions similar to (3a) and (3b) and only on one ungrammatical condition as (3c), so interference effects have not been systematically examined especially across ungrammatical conditions.

Grammatical insensitivity to number agreement was also reported in Chen, Shu, Liu, Zhao, and Li's ERP study (2007). Their study mainly aimed to examine whether L2 processing of S-V agreement might be influenced by L1 specific properties. Chinese L2 speakers of English, whose native language lacks S-V agreement morphosyntactic features, were presented with sentences like (3a-d) and asked to judge their grammaticality while their ERP responses were recorded during reading. The results showed that L2 speakers were highly accurate in their grammaticality judgements. However, their ERP patterns were significantly distinct from L1 speakers' in not exhibiting any grammaticality effects. No evident similarity-based interference effects were observed across grammatical or ungrammatical conditions for both groups.

In another ERP study, Tanner, Nicol, Herschensohn and Osterhout (2012) examined interference effects in S-V agreement processing in native English speakers and Spanish L2 speakers of English. Grammaticality, which is based on the verb's agreement with the singular

head, the structure of the subject NP and the distractor number were all manipulated as shown in (30).

(30) a. The winner of the big trophy/trophies has/*have proud parents.

b. The winner who got the trophy/trophies has/*have proud parents.

Unlike Jiang (2004) and Chen et al. (2007), the results showed a significant main effect of grammaticality for both groups. The P600 effects on the verb, when it disagreed with its subject, were smaller in Spanish L2 speakers of English. However, both groups showed smaller effects when the verb followed a plural distractor compared to a singular distractor, indicating interference effects. Though P600 effects were larger when the distractor was embedded in a relative clause as in (30b) compared to when it is in a prepositional phrase as in (30a), there was no significant interaction between the distractor number, subject NP structure and group. This suggests that interference effects were not modulated by the subject NP structure, and both groups displayed similar interference profiles. Interference effects, however, were not observed in grammatical conditions. The results could be taken as evidence of the role of interference during agreement processing rather than shallow parsing. That is, if the latter is the case, as pointed out by Tanner and colleagues, and that L2 speakers' reduced grammaticality effects caused by the plural distractor in ('*The winner of the big trophies have …') was seen as a result of shallow parsing, grammatical sentences with a plural distractor as in 'The winner of the big trophies has…' would also be perceived as ungrammatical. However, this was not the case. The study indicated that processing differences between L1 and L2 speakers were not qualitative, but quantitative in nature.

Based on the studies above, one may infer that L2 speakers show native-like S-V agreement processing only if the same structure is also instantiated in the L1 since Spanish has S-V agreement while Chinese does not. However, recent studies have found that L2 speakers could parse S-V agreement similarly to L1 speakers even if their L1 lacks overt agreement morphology as English (Cheng, Cunnings, Miller, & Rothman, 2021; Cheng, Rothman, & Cunnings, 2022; Lim & Christianson, 2014). For instance, Korean L2 speakers of English with varying proficiency were examined by Lim and Christianson (2014) for S-V agreement, which is absent in their L1, using eye-tracking during reading. They used sentences like (31) in which the distractor inside the relative clause was manipulated in number across grammatical and ungrammatical conditions.

(31) The teacher(s) who instructed the student(s) were very strict.

During real-time reading, L1 speakers showed clear grammaticality effects at the verb and spillover region, and similar effects were demonstrated by L2 speakers at the spillover region. Although sensitivity to grammaticality in L1 speakers was consistently found in most reading time measures compared to L2 speakers, the sensitivity of L2 speakers was observed to be affected by proficiency, becoming more evident as proficiency increased. In response to the distractor, the ungrammaticality effects were attenuated for both groups in conditions with a matching plural distractor indicating facilitatory interference. The results of this study revealed that L2 speakers can eventually detect agreement violations regardless of their L1, and proficiency can considerably modulate L2 processing in line with other previous agreement studies (Coughlin & Tremblay, 2012; Hopp, 2006; Jackson, 2008; Sagarra & Herschensohn, 2010).

More recently, Lee and Philips (2022) reported that advanced Korean L2 speakers could even outperform L1 speakers in a speeded judgement task. According to their findings, both groups' acceptance rates showed facilitatory interference in ungrammatical sentences containing relative clauses (e.g., '*The artist who made the sculpture/s are very talented.'). However, L2 speakers did not show any interference effects in sentences containing prepositional phrases (e.g., '*The artist with the tall sculpture/s are very talented.'), while L1 speakers did. Lee and Phillips argued that L2 speakers may utilize an additional monitoring mechanism that filters out ungrammatical structures compared to L1 speakers, at least when making explicit judgements. Similar effects were absent from both groups' reaction times, however.

Taken together, although earlier studies suggested L2 speakers' insensitivity to agreement violations, as predicted by the SSH, more recent research suggest that L1 and L2 speakers can behave similarly in the computation of S-V agreement.

2.6.2 L2 processing of reflexive-antecedent dependency.

Though some studies in adult L2 acquisition research examined reflexive dependencies using off-line experiments to address issues related to L2 interlanguage grammar from Chomsky's (1981) view of principles and parameters (see Hawkins, 2001; White, 2003 for a review), fewer studies investigated this phenomenon from a processing perspective (Felser & Cunnings, 2012; Felser et al., 2009).

Felser et al. (2009), for instance, aimed to examine the degree to which Japanese L2 speakers of English are sensitive to syntactic constraints on reflexives relative to native English speakers. In an offline antecedent choice task, both groups performed highly accurately, indicating that L2 speakers could use the locality and c-commanding conditions for reflexive binding. In an eye-tracking experiment, only grammatical sentences were included and manipulated by two factors: the gender match between the distractor and the reflexive and the

syntactic structure so that the distractor either c-commanded the reflexive or did not, as illustrated below.

(32) John/Jane and Richard were very worried in the kitchen of the expensive restaurant.

a. c-command

[John/Jane] noticed that Richard had cut himself with a very sharp knife.

b. No c-command

It was clear to [John/Jane] that Richard had cut himself with a very sharp knife.

The results showed that L2 speakers took longer first-pass reading times to read a reflexive when a c-commanding but non-local distractor matched the reflexive's gender than when it did not. This can exhibit inhibitory interference, but it was not interpreted on this basis. No gender match effects, however, were found for the non-c-commanding distractor. These findings were consistent with the results of an additional timed grammaticality judgement task in which L2 speakers were less accurate than L1ers at judging sentences when the matching illicit distractor c-commanded the reflexive. As there were no reliable effects of the distractor gender match in any later regions or measures for both groups, it was suggested that L2 speakers' ultimate antecedent interpretations were native-like. Felser et al. (2009) concluded that L2 speakers' native-like performance in the untimed task showed their reliance on grammatical knowledge. However, their reading times revealed that they were initially influenced by pragmatic information. The presence of a matching discourse-prominent distractor was argued to have caused processing difficulty for L2 speakers, though L1 influence cannot also be totally precluded since Japanese allows long-distance reflexive binding by a non-local antecedent. The results, in

fact, might alternatively be attributed to similarity-based interference effects, as the increased reading times observed at the reflexive region can result from the overlap in gender feature between the target and distractor, causing that processing difficulty. If L2 speakers were really influenced by the discourse-prominent distractor as argued, we would also expect to observe an early effect in the condition where it did not match the reflexive's gender as it contradicted their expectation, but this was not the case. The role of memory retrieval interference in the L2 processing of reflexives, therefore, clearly requires more investigation.

Building on the study of Felser et al. (2009) and to exclude the possibility that Japanese L2 speakers' processing was influenced by L1 transfer of long-distance binding, Felser and Cunnings (2012) investigated whether German L2 speakers of English, whose L1 is similar to English in that a reflexive can only be bound by a local antecedent, would be affected by a non-local distractor during processing. Their first eye-tracking experiment's design included conditions similar to Sturt's (2003) Experiment 1, as exemplified by (33).

(33) **a. Target-match/Distractor-mis/match**

Helen/James has worked at the army hospital for years. She/He noticed that the soldier had wounded himself while on duty in the Far East. Life must be difficult when you are in the army.

b. Target-mismatch/Distractor-mis/match

James/Helen has worked at the army hospital for years. He/She noticed that the soldier had wounded herself while on duty in the Far East. Life must be difficult when you are in the army.

Though their responses in an offline antecedent choice task confirmed their knowledge of binding condition A, eye-tracking data revealed that German L2 speakers initially displayed a main effect of the distractor gender whereas effects of the target could only be observed in later processing measures, as opposed to native English speakers who only displayed a significant main effect of the target in all measures. The results supported those of Felser et al. (2009), but German L2 speakers showed a slowdown in the reflexive when the distractor mismatched the reflexive while Japanese showed it when the distractor matched it.

Felser and Cunnings (2012) conducted a follow-up experiment using materials similar to those in Experiment 1, but with the distractor not c-commanding the reflexive as in (e.g., 'The soldier that he/she treated on the ward wounded himself/herself…') to determine whether L2 speakers' coreference assignment in Experiment 1 was influenced by the distractor's discourse prominence or by their preference to link the reflexive to the matrix subject, which might be argued that it involves a binding requirement since it c-commands the reflexive. Despite the fact that the matrix subject in Experiment 2 was the target antecedent, German L2 speakers initially focused on the non c-commanding embedded distractor. The native group also experienced some initial distractor effects, spending longer reading times when it mismatched the reflexive than when it did not, but their reading times were also influenced by the gender of the target. The results were taken to argue that L2 speakers were initially influenced by the discourse-prominent distractor antecedent regardless of whether it c-commanded the reflexive or not. Given that German allows only local binding, L1 transfer was excluded as a cause for L2 speakers' early preference for the non-local distractor. Instead, Felser and Cunnings argued that L2 speakers had difficulty applying binding constraints.

This initial preference for a discourse prominent antecedent may be compatible with the SSH (Clahsen & Felser, 2006a, 2018). Cunnings (2017) alternatively argued that L2 speakers may construct fully specified syntactic parses but may be more prone to interference than L1

speakers, if they weigh non-syntactic retrieval cues, especially those related to discourse prominence, more highly than L1 speakers.

2.7 The Present Study

2.7.1 Gaps and research aims.

In non-native processing, the literature on processing discontinuous dependencies revealed somewhat mixed findings. Some studies suggest L1 and L2 speakers resolve dependencies in a similar manner, while other suggest L2 speakers do not compute linguistic dependencies in a target-like way (see Dallas & Kaan, 2008; Felser, 2015, 2019; Roberts, 2013 for a review). L2 speakers' performance may also vary from one kind of dependency to another, as in the S-V agreement and reflexive dependencies reviewed above (Felser et al., 2009; Felser & Cunnings, 2012; Jiang, 2004; Tanner et al., 2012). Furthermore, possible L1 effects are sometimes not found in L2 speakers' processing patterns, but L2 processing may also be non target-like even if the L2 shares the same properties of the phenomenon being investigated with the L1 (Felser, 2015). The use of different methods, different L1/L2 combinations, and variations between participants and contexts may additionally cause conflicting findings across studies.

In addition to research findings variability, various theoretical accounts have been proposed to explain differences in L1/L2 processing. Some argue that L2 processing is not syntactically driven and less attention is paid to grammatical information relative to L1 processing (Clahsen & Felser, 2006a, 2018). Others, on the other hand, contend that L2 processing can be influenced by individual differences while it is fundamentally not different from L1 processing (Hopp, 2014; Lim & Christianson, 2014; McDonald, 2006; Sagarra & Herschensohn, 2010). More recently, Cunnings (2017) argue that L2 speakers construct similar syntactic parses, but they could be more susceptible to interference effects that may arise from the cue-based memory retrieval mechanism. L2 processing has been investigated thoroughly in the literature in terms of shallow parsing and working memory capacity imitations. However, the working memory mechanisms that underpin language comprehension are not fully clear, and little is currently known about whether differences in retrieving information established during real-time

processing from memory can better explain potential L1/L2 processing differences (Cunnings, 2017). Additionally, the nature of memory retrieval mechanism across dependencies, particularly between S-V agreement and reflexives, which has been debated in L1 processing has not also been examined in L2 processing.

The aim of this research, therefore, is to tease apart the different accounts of dependency resolution during L2 processing and to elucidate the contrasting existing L2 findings between agreement and reflexives. Previous L2 studies suggest native-like processing of S-V agreement is possible for L2 speakers, while L1/L2 differences have been observed during reflexive resolution. However, existing studies have examined either S-V agreement or reflexive processing, making it difficult to compare the L2 processing patterns of these two dependencies. Studies have also used different experimental designs and methodologies, besides testing different L2 groups. The current study aims to control for confounding factors that may have created discrepancies in L2 processing research findings by investigating the processing of S-V agreement and reflexive-antecedent dependencies in the same large-scale group of Arabic L2 speakers of English. Given that the two phenomena share similar properties in Arabic and English, L1 influence will be minimised, and a greater focus will be placed on the mechanisms underlying L2 processing. It is also the purpose of this study to take some individual differences into account, such as L2 proficiency and lexical automaticity for their potential impact on parser performance. In view of the foregoing, the findings of such a well-controlled study will offer a better understanding of the nature of the mechanisms employed by L2 speakers during processing linguistic dependencies and the factors that may influence their real-time sentence comprehension.

2.7.2 Research questions.

To fulfil the research aims, the subsequent research questions will be addressed, followed by various predictions regarding L2 language processing based on different theoretical perspectives.

1. Does Arabic L2 speakers' language processing differ from native English speakers' in sentence comprehension?

2. To what extent can the cue-based memory retrieval mechanism influence L2 processing?

3. Is Arabic speakers' L2 processing of S-V agreement and reflexive-antecedent dependencies influenced by their L2 proficiency level?

4. Does lexical automaticity influence S-V agreement and reflexive processing in Arabic and English speakers?

5. Are S-V agreement and reflexive-antecedent dependencies uniformly processed?

2.7.3 Predictions and hypotheses.

2.7.3.1 Shallow Structure Hypothesis

This generally predicts that L2 processing differs from native processing due to reduced sensitivity to structural information, such that a given grammatical constraint might not be applied in a native-like way. L2 processing is therefore expected to rely heavily on lexical semantics, plausibility, and surface-level information more than grammatical constraints during real-time processing. In other words, it hypothesizes that L2 processing is primarily driven by semantic or discourse-related cues rather than syntactic ones, which means that the former cues are efficiently fast utilised and prioritised along with other types of non-grammatical information over the latter ones compared to L1 parsers (Clahsen & Felser, 2018, p. 695). A further contention in the SSH is that L2 speakers from different L1 backgrounds would similarly underuse grammatical information and adopt shallow processing regardless of their L1. It was also speculated that L2 processing difficulties might be specifically observed in discontinuous syntactically-mediated dependencies in general (Clahsen & Felser, 2006a), and in backward-looking dependencies in particular, which requires reaccessing previously established representations, such as reflexive-antecedent dependencies, compared to forward-looking dependencies (Felser, 2015). That is, it has been argued that L2 speakers' reduced ability to build detailed hierarchical sentence representations in the first place and/or the instability of the syntactic representations constructed during L2 processing, compared to L1 processing, make backward-looking dependencies more challenging (Felser, 2015).

Accordingly, with regards to the dependencies under investigation, Arabic L2 speakers' performance will be compatible with the Shallow Structure Hypothesis if their reading times are not guided by structural constraints, such as hierarchical syntactic structure and binding Principle A, but instead are influenced more by structurally illicit feature-matching distractors and linear parsing heuristics. This will be evidenced generally by being significantly different from L1

speakers in not showing grammaticality effects, or in accepting ungrammatical sentences with matching distractors and equivalently rejecting grammatical sentences with mismatching distractors. Moreover, L2 speakers might find processing a reflexive-antecedent dependency more challenging than a subject-verb agreement dependency since it is more clearly a backward-looking dependency. S-V agreement processing might be seen by some as a forward-looking dependency if the process involved is believed to be only predictive (Tanner, Nicol, & Brehm, 2014). However, it has been argued that it may also be a hybrid type of dependency since it may involve both a prediction mechanism (forward-looking search) and a cue-based retrieval mechanism (backward-looking search) when a verb's number cue is used to locate the target subject in memory (Dillon et al., 2013; Felser, 2015; Wagers et al., 2009). The case is different for reflexives since they cannot be anticipated. Therefore, a reflexive-antecedent dependency is reliably a backward-looking dependency, which is predicted to be more problematic.

2.7.3.2 Cue-Based Memory Retrieval Interference Account

The assumption that linguistic dependencies processing relies on a cue-based retrieval mechanism entails parsers' potential susceptibility to interference effects. The memory retrieval interference account predicts that the presence of similar items in memory may adversely influence dependency resolution processing. Items that share similar features with the target dependent element that needs to be retrieved for dependency resolution may inhibit the smooth retrieval of the target in grammatical sentences causing reading times to increase. However, they may also facilitate processing ungrammatical sentences and speed-up reading times when a target dependent element is not available. As far as L2 processing is concerned, Cunnings (2017) argued that L2 speakers can employ syntactic or structure-building parsing strategies, but they might be more prone to interference effects relative to L1 speakers. This means that L2 speakers are expected to be sensitive to grammatical constraints, but their performance might be more attenuated by associated retrieval effects.

Accordingly, L2 speakers are predicted to show grammaticality effects as L1 speakers, but they may spend longer reading times in grammatical conditions that particularly involve matching distractors than those that do not, and shorter reading times in ungrammatical conditions with matching distractors relative to those with mismatching distractors. Such interference effects are expected to be more pronounced in L2 speakers in both S-V agreement and reflexive-antecedent dependencies, compared to L1 speakers who have usually shown facilitatory interference specifically in S-V agreement dependencies.

2.7.3.3 Individual Differences Factors

Individual differences in factors, such as proficiency or lexical automaticity as a cognitive or capacity-related aspect of language processing, among other factors, have been argued to play a significant role in the course and outcomes of non-native syntactic processing (Hopp, 2006, 2010, 2014, 2015, 2018).

Based on previous studies (e.g., Coughlin & Tremblay, 2012; Hopp, 2006; Jackson, 2008; Rossi et al., 2006; Sagarra & Herschensohn, 2013; Steinhauer, White, & Drury, 2009), advances in proficiency have been suggested to positively correlate with the growing use of morphosyntactic information and target-like sentence processing. That is, success in integrating different information types, such as lexical-semantic and morpho-syntactic information, in parallel during real-time processing is likely to be associated with proficiency since proficiency generally represents the mastery level of an L2 (Hopp, 2015). Individual differences in proficiency thus may predicate whether L2 speakers recruit different grammatical information and whether they ultimately achieve native-like processing.

The automaticity of lexical processing has also been argued to interact with structure-building operations and syntactic processing (The Lexical Bottleneck Hypothesis, Hopp, 2018). As lexical automaticity reflects the dynamics of lexical access and retrieval that feed into syntactic

processing, less automated individuals with slower lexical processing are assumed to deplete resources that could subserve syntactic processing. Such higher demands on lexical processing may ultimately leave them with insufficient resources to complete target-like syntactic parses. Accordingly, differences in lexical automaticity between individuals were suggested to engender differences in syntactic processing (Hopp, 2018).

Taken together, individual differences in the aforementioned factors are expected to moderate sensitivity to grammaticality and influence processing patterns. Specifically, higher L2 speakers in proficiency are predicted to show more pronounced grammaticality effects and target-like syntactic processing. Similarly, L2 speakers with higher degrees of lexical automaticity are expected to show greater grammaticality effects and native-like processing patterns.

CHAPTER THREE

3. Study 1

3.1 Introduction

This study aimed to examine L1 Arabic speakers' L2 processing of S-V agreement and reflexive-antecedent dependencies in English to uncover the nature of the processing mechanism involved by L2 speakers during dependency resolution compared to L1 speakers and the factors that may influence their processing patterns. To gain a better understanding of the discrepancies between native and non-native speakers' performance, all participants completed an offline and online task. While the offline grammaticality judgement task will tap into their grammatical knowledge, the online self-paced reading task can reveal areas of processing difficulties that might not be clear from their performance in the offline task, and provide an assumption of the mechanisms underlying their sentence processing. A lexical decision task for both groups along with an L2 proficiency test for L2 speakers were also used to examine how individual differences in lexical automaticity and proficiency may affect sentence processing.

3.2 Method

3.2.1 Participants.

A total of 188 Arabic L2 English speakers and 189 L1 English speakers were recruited. They were recruited from the University of Reading student community or by advertising on different media platforms. They were contacted then via email. Participants took part voluntarily or received course credit for their participation.

Based on the background information questionnaire they filled out (see Appendix A), only Arabic L2 English speakers who started learning English at age 5 or after were included. This led to the removal of eight participants. For L1 English speakers, only those who identified English

as their only native language and considered themselves not bilingual, meaning that they did not have a native-like command of languages other than English, were included. This led to the removal of ten participants. An additional two L2 speakers were also removed for not completing the second part of the study, and another two L1 speakers were also removed; one for not pressing the right buttons in the grammaticality judgement task and the other for not completing the second part. We also excluded two L2 speakers and one L1 speaker from the data analysis, as clarified in the next section, due to their unusually fast median reaction times in the grammaticality judgement task, which indicated a lack of attention.

Accordingly, 176 L2 English speakers (44 males, mean age = 30, range = 18-43) and 176 L1 English speakers (37 males, mean age = 28, range = 18-62) could be included. 148 L1 English speakers were British, 21 were American, three were Australian, three were Canadian and one was New Zealander whereas all Arabic L2 English speakers were from Saudi Arabia. The L1 group consist of 17 PhD students, 37 Master's students, 86 Bachelor students, alongside 36 individuals who did not have student status. The majority of the L2 speakers in this study were also students, compromising 96 PhD students, 36 Master's students, 10 Bachelor students and 19 English course students, while 15 participants were not affiliated with any educational institution during the time of testing. All L1 speakers were exclusive native speakers of English, and all L2 speakers were exclusive native speakers of Arabic, with neither group having a native-like command of any additional languages. L2 speakers reported that they commenced their English language learning in a school setting at an average age of 13 years, with a range spanning from 5 to 36 years and a median of 13 years.[11] They were either in an English-speaking country during

[11] 7 participants started their English language learning at age 5, 13 at age 6, 6 at age 7, 6 at age 8, 4 at age 9, 14 at age 10, 7 at age 11, 28 at age 12, 33 at age 13, 8 at age 14, 8 at age 15, 5 at age 16, 1 at age 17, 4 at age 18, 3 at age 19, 4 at age 20, 2 at age 22, 5 at age 23, 6 at age 24, 2 at age 25, 2 at age 27, 3 at age 28, 2 at age 31, 1 at age 33, 1 at age 35 and 1 at age 36.

testing or they had been in one previously[12]. Their immersion in an English-speaking country ranged from 1 to 180 months, with a mean immersion duration of 44 months. Notably, only nine of them were immersed before the age of 18, while the rest experienced immersion after that age, with an average age of 27. The majority reported that they used Arabic at home with their families, while only 15% reported using both Arabic and English. The daily usage of English varied among individuals, ranging from 0 to 12 hours, with an average of 3 hours.[13] Except for 10 participants, the majority had undertaken either the IELTS or TOEFL examination, and their most recent scores fell within the CEFR levels of B1 and C2. L2 English speakers' current proficiency was additionally assessed through the Oxford Quick Placement Test, resulting in an average score of 42/60 (SD = 9.7, range = 22-60) with the majority could be classified as intermediate to advanced learners.[14] The correlation coefficient value between their IELTS/TOEFL scores and the Oxford Quick Placement Test is 0.74, signifying a robust positive correlation.

3.2.2 Materials.

3.2.2.1 Grammaticality Judgement Task (GJT).

This task aimed to assess L2 speakers' grammatical knowledge of agreement and reflexives, and their ability to discriminate between grammatical and ungrammatical sentences. Stimuli consisted of 24 sentences testing S-V agreement as in (34) and 24 sentences testing reflexives as in (35), such that each participant was shown six different items of each of the four conditions for each dependency type. For both dependencies, grammaticality effects were tested in two baseline 'no distractor conditions' as in (34a/b) and (35a/b) where the verb mismatched its

[12] The clarity and reliability of the effects resulting from naturalistic exposure or immersion on their performance can be uncertain due to differences in the timing of immersion experiences in comparison to the time of testing across participants.

[13] While the question focused on the number of hours spent outside the classroom or work environment rather than inside, a few respondents may have included hours from both settings in their responses, making the information somewhat uncertain.

[14] 49 participants scored from 50 to 60, 59 participants scored from 40 to 49, 44 participants scored from 30 to 39, and 24 participants scored from 20 to 29.

subject in number in S-V agreement dependencies and the reflexive mismatched its antecedent in gender in reflexive dependencies to confirm their basic understanding that subjects and verbs must agree in number, and that reflexives require a gender-matching antecedent. Grammaticality effects were also examined in another two structurally parallel sentences to the baseline conditions but with a distractor as in (34c/d) and (35c/d) to assess L2 understanding of what constituents must have the same number/gender as the verb or reflexive. The distractor in the 'distractor conditions' was consistently embedded in a prepositional phrase, which matches the syntactic context of the distractor in the online task.

(34) S-V Agreement

a. *No Distractor, Grammatical*

The boys were hurt yesterday afternoon.

b. *No Distractor, Ungrammatical*

*The boy were hurt yesterday afternoon.

c. *Distractor, Grammatical*

The boys near the girl were hurt yesterday afternoon.

d. *Distractor, Ungrammatical*

*The boy near the girls were hurt yesterday afternoon.

(35) Reflexives

a. *No Distractor, Grammatical*

The man cut himself two hours ago.

b. *No Distractor, Ungrammatical*

*The man cut herself two hours ago.

c. *Distractor, Grammatical*

The man behind the lady cut himself two hours ago.

d. *Distractor, Ungrammatical*

*The man behind the lady cut herself two hours ago.

For agreement dependencies, the verb was plural and held constant across all conditions as in the online task. Hence, grammaticality was based on the plural verb's agreement with its target subject in number. The distractor in the distractor condition mismatched the subject and the verb in number in the grammatical condition, but it matched the verb in the ungrammatical condition. Accordingly, the distractor's presence would further show the extent to which participants have solid grammatical knowledge of permissible and impermissible conditions.

For reflexive dependencies, the subject-gendered noun phrase was singular and kept constant across all conditions. Both the subject and the reflexive were located within the same local clause. The subject was a male-biased noun in half of the experimental items and a female-biased noun in the other half. Grammaticality was based on the reflexive's match with the subject, its target c-commanding antecedent, in gender. Similar to agreement dependencies, the non c-commanding distractor in the distractor condition mismatched the subject and the reflexive in gender in the grammatical condition. It, however, matched the reflexive in the ungrammatical condition. Participants' judgement thus would indicate whether it is guided by the binding Principle A condition or is affected by the distractor's presence.

The 48 experimental items were distributed across 4 lists in a Latin square design and were mixed with 64 filler items of different sentences structures, with half being grammatical and half ungrammatical. The complete set of stimuli can be found in Appendix B.

3.2.2.2 Self-Pased Reading Task (SPR).

To examine real-time sensitivity to grammaticality and vulnerability to interference during S-V agreement and reflexive processing, 24 item sets of 4 conditions were constructed like those in (36) and (37) for each dependency type, with grammaticality (grammatical vs. ungrammatical) and distractor match (distractor match vs. distractor mismatch) as factors. Grammaticality was manipulated by varying the number feature of the subject in S-V agreement dependencies and the gender feature of the subject in reflexive-antecedent dependencies. Likewise, the second factor was manipulated by varying the number of the distractor in S-V agreement dependencies and the gender of the distractor in reflexive-antecedent dependencies making it either match or mismatch the critical region (i.e., the verb and the reflexive region where retrieval of the target subject is triggered in each dependency) across the grammatical and ungrammatical conditions.

(36) S-V Agreement

a. *Grammatical, Distractor Match*

   The waitresses near the schoolgirls unsurprisingly were unhappy about all the noise.

b. *Grammatical, Distractor Mismatch*

   The waitresses near the schoolgirl unsurprisingly were unhappy about all the noise.

c. *Ungrammatical, Distractor Match*

   *The waitress near the schoolgirls unsurprisingly were unhappy about all the noise.

d. *Ungrammatical, Distractor Mismatch*

*The waitress near the schoolgirl unsurprisingly were unhappy about all the noise.

(37) Reflexives

a. *Grammatical, Distractor Match*

The man near the policeman strangely isolated himself from society for many years.

b. *Grammatical, Distractor Mismatch*

The man near the policewoman strangely isolated himself from society for many years.

c. *Ungrammatical, Distractor Match*

*The lady near the policeman strangely isolated himself from society for many years.

d. *Ungrammatical, Distractor Mismatch*

*The lady near the policewoman strangely isolated himself from society for many years.

In both dependencies, the subject and the distractor were animate nouns to ensure that neither of them provides additional cues beyond the contribution of syntactic position and number in agreement conditions or gender in reflexive conditions. The subject was modified by a prepositional phrase in which the distractor was embedded. Testing interference effects in reflexive dependencies using this syntactic context for the distractor has not been examined before. As interference has not usually been observed in reflexive dependencies, the processing of reflexives might be more vulnerable to interference from distractors in PPs, because they are

less distinctive in memory as proposed by Van Dyke and McElree (2011), compared to distractors in core argument positions (i.e., subject or object), which have been examined in previous studies. This has been shown in subject-verb thematic binding (Van Dyke & McElree, 2011) and S-V agreement (Parker & An, 2018). To avoid any potential spillover processing effects from the distractor on the verb region, an adverbial was placed after the distractor and before the main verb. This was done because plural nouns generally take longer to process than singular nouns, and this effect could potentially affect the processing of the following words (Wagers et al., 2009). The adverb thus provides 'padding' to deal with this spillover before the critical verb. The critical region (i.e., the verb in agreement conditions and the reflexive in reflexive conditions) was then followed by five words.

For agreement conditions, the main verb was always a plural verb *be* in the past tense 'were' and held constant across all conditions. While some studies have used singular and plural verbs in grammatical and ungrammatical conditions respectively (e.g., Wagers et al., 2009), we kept the verb constant across conditions to avoid any potential differences in reading times at the critical region merely due to differences in the length of the verb itself. Grammatical sensitivity to agreement violations should elicit significant grammaticality effects from spending longer reading times in (36c/d) than in (36a/b). However, if L2 speakers employ shallow processing (Clahsen & Felser, 2006a, 2018), their reading times should reveal linear parsing heuristics or distractor effects in both (36b) and (36c), such that (36b) should be read slower than (36a) and (36c) should be read faster than (36d). The cue-based memory retrieval model (Lewis & Vasishth, 2005), on the other hand, predicts that resolving dependencies can be influenced by two types of interference effects. In our grammatical conditions, the subject and the verb were plural agreeing in number. Inhibitory interference is predicted when the distractor is also plural matching the verb in number as in (36a), leading to longer reading times in (36a) than in (36b). In the ungrammatical conditions, however, facilitatory interference is predicted when the distractor matches the plural

verb in the absence of a grammatically matching subject, and so (36c) should be read faster than (36d). L2 speakers, in particular, are predicted to show increased interference effects (Cunnings, 2017).

For reflexive conditions, singular gender-biased nouns (e.g., schoolboy, businesswoman) were used for the target subject and distractor rather than gender-stereotyped nouns (e.g., surgeon) to avoid any potential cultural differences in the stereotypical gender they refer to or the possibility that they can eventually be taken to refer to both genders. Half of the items contained the reflexive *himself* and the other half contained *herself*. Though both the target antecedent and the distractor are within the reflexive's local domain, 'the man' in (37a), for instance, is the only acceptable antecedent for the reflexive since it c-commands it unlike 'the policeman' that does not. If binding Principle A is respected, grammaticality effects manifested by longer reading times in (37c/d) than in (37a/b) should be observed. The Shallow Structure Hypothesis (Clahsen & Felser, 2006a, 2018) predicts that L2 speakers might instead link the reflexive to the linearly closest antecedent or be influenced by the non-c-commanding antecedent, and so reading times should be slower in (37b) than in (37a) and (37c) should also be read faster than (37d). The cue-based model (Lewis & Vasishth, 2005), alternatively, predicts longer reading times in (37a) than in (3b) and shorter reading times in (37c) than in (37d) as a result of inhibitory and facilitatory interference effects in (37a) and (37c) respectively from gender-matching distractors. Cunnings (2017) predicts that such interference effects would influence L2 speakers more than L1 speakers.

In addition to the 48 experimental items, 64 structurally different fillers were combined in a pseudo-randomised Latin-square design. To prevent participants from subconsciously recognising the distribution of stimuli, the fillers included items with different types of anaphors in positions that were distinct from the reflexive's position in the experimental items. Some fillers also contained main verbs that do not require agreement morphology, or verbs that differed in

form or number from those in the S-V agreement experimental items (see Appendix C). To ensure that participants read for meaning, all the experimental items and fillers were followed by a yes/no comprehension question, half of which are answered by yes and the other half by no. The question never asked about any aspect of the relationship between the verb and its subject or the reflexive and its antecedent to prevent any cues that may make participants aware of the experiment's manipulation.

3.2.2.3 Lexical Decision Task (LDT).

Following Hopp (2014), an adopted lexical decision task was used to examine the extent to which individual differences in lexical automaticity may affect syntactic processing. Forty words, originally taken from the most frequently generated English items in the category norms list in Van Overschelde, Rawson, & Dunlosky (2004), were included. Another forty English pseudowords were created by mimicking other frequent semantically-related items but with a one-syllable change considering English phonotactic constraints (e.g., nail → nait). Those items were presented in the centre of the screen successively in random order, and participants should decide within 5000 ms whether a word is real or not. Accuracy and response latency were recorded. To compute automaticity, the coefficient of variance (CV) was used as a measure, defined as the standard deviation of response times for each participant divided by the average response times of the words judged correctly (Hopp, 2014; see Segalowitz & Segalowitz, 1993 for a discussion). CV signifies the extent of variability in word recognition (latency variability) relative to total performance time. If the variance is low, it indicates fewer effortful processes are used and processing is automatic.

3.2.3 Pilot testing.

A pilot study was conducted on 35 American and British native English speakers (15 male, mean age=41, range=21-76), none of whom took part in the main experiment, to test the stimulus

materials, the feasibility of tasks, the randomisation procedure, and participant recruitment and retention. Given that both L1 and L2 speakers would undergo the same tasks[15] and the fact that it might be difficult to recruit a sufficient number of L2 participants later for the main study, only native speakers were included in the pilot testing. They were contacted by email for two sessions, a week apart. The preliminary investigation showed generally a consistent range of responses in comprehension accuracy and grammaticality judgements validating target responses. A few items that had error rates exceeding 14%, which means more than 5 participants answered or judged them differently, were revised and amended as follows. Two comprehension questions were changed due to low comprehension accuracy compared to other questions, as each was wrongly answered by around 20% and 31% of participants. Two filler items in the GJT were also reworded because they were incorrectly judged by around 26% and 54% of participants. Moreover, due to some native speakers' feedback suggesting that the noun ‘actor’ may refer to both males and females and their variations in responses accordingly, it was replaced with other clearer gender-biased nouns in all reflexive items.

In terms of presentation, the code relevant to the order of items was also revised and tested again to ensure not only that items from the same condition do not appear next to each other, but also that experimental items generally do not appear consecutively.

3.2.4 Procedure.

All tasks were conducted using the web-based IbexFarm software (Drummond, 2013). This has been shown to yield reliable results when used in previous psycholinguistic studies (Chemla, Cummins, & Singh, 2017; Dillon, Clifton, & Frazier, 2013; Lago et al., 2015; Parker & An, 2018; Parker & Phillips, 2016; Wagers & Phillips, 2014). The study was completed over two

[15] Only the proficiency test will additionally be administered to L2 speakers, but this has been extensively used in previous studies.

sessions with at least one week apart. Each participant was contacted by email and given a unique reference number to use in the two sessions.

In the first session, participants filled out a background questionnaire after reading the information page and consenting. They then completed the self-paced reading experiment. In this experiment, each sentence was initially presented as a series of dashes that masked the sentence's words, and participants needed to press the space bar to reveal each word. The presentation was non-cumulative, meaning that the previous word was hidden by a dash once the next word appeared. After reading each sentence, a counterbalanced yes/no comprehension question appeared on a separate screen. Participants needed to answer by pressing the "1" key for yes or "2" for no. Prior to the test items, participants completed three practice items to become familiar with the task. After the self-paced reading experiment, L2 speakers completed the Oxford Quick Placement Test. The first session took approximately 30-40 minutes for L1 speakers and 50-60 minutes for L2 speakers.

In the second session, participants completed the grammaticality judgement experiment in which they were required to judge the grammaticality of the sentences based on their intuitions. In this experiment, each sentence was presented one at a time in its entity by pressing the space bar. Below each sentence, two choices were given (grammatical/ungrammatical) to which participants responded by pressing the "1" key if they think the sentence is grammatical or "2" if they think it is not. Two practice items preceded the test items, and there was no time limit to finish the task. Participants afterward completed the lexical decision task. The second session took around 30 minutes for native speakers and 30-40 minutes for L2 speakers.

3.2.5 Data analysis.

Before data analysis, accuracy on comprehension questions in the self-paced reading experiment was checked to include data only from participants with at least 75% accuracy. All participants scored over that threshold value, as an index that they paid attention. Mean comprehension accuracy rates across the experimental and filler items were 96% for L1 speakers (range= 86-100%) and 94% for L2 speakers (range= 78-100%). Additionally, the median reaction time for each participant across the experimental and filler items in the grammaticality judgement experiment was checked as a measure of attention. Only participants with a median reaction time greater than 1500 ms were included in the analysis. This resulted in the removal of the two L2 speakers and one L1 speaker mentioned earlier. Moreover, reading times shorter than 100 ms or longer than 10,000 ms in the self-paced reading were excluded, since these likely reflect unintentional button presses or loss of attention. This affected 0.01% and 0.02% of L1 and L2 speakers' reflexives data, respectively, and 0.02% and 0.04% of S-V agreement data for L1 and L2 speakers.

Analysis was carried out separately for each dependency type using R (version 4.0.3) (R Development Core Team, 2020). Grammaticality judgement accuracy data with a binomial distribution were analysed using generalised mixed-effects models with the lme4 package (Baayen, Davidson, & Bates, 2008; Bates, Mächler, Bolker, & Walker, 2015). The analysis included accuracy (0=ungrammatical, 1=grammatical) as the dependent variable with by-subject and by-item random effects, and sum-coded (-0.5/0.5) main effects of group (L1/L2), grammaticality (grammatical/ungrammatical), distractor's presence (distractor/no distractor) and their interactions as fixed effects. In all statistical models reported below, we initially fit models with the maximal random effects structure including by-subject and by-item random intercepts, and random slopes for each fixed effect (Barr, Levy, Scheepers, & Tily, 2013). If the maximal model failed to converge, we refit the model after removing the correlation parameters. If this

model still did not converge, we iteratively removed the random effects that accounted for the least variance until convergence was achieved. If an interaction was observed, follow-up analysis was performed using nested contrasts. The *p* values for each fixed effect were measured using the Satterthwaite approximation by the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017).

For the self-paced reading experiment, reading times were log-transformed to remove skewness. Analysis was performed for reading times at the critical region ('himself'/'herself' in reflexive dependencies and 'were' in S-V agreement dependencies) and spillover region (the word that directly followed the critical region) following a similar procedure as with the grammaticality judgement data, but linear mixed effects models were used instead. Fixed effects included sum-coded main effects of group (L1/L2), grammaticality (grammatical/ungrammatical), distractor match (match/mismatch) and their interactions. In addition, region (critical/spillover) was included as a fixed effect to reduce the number of analyses conducted on each region and to account for any potential effects that may result from the time course of reading processes across regions. This can increase the reliability of findings rather than drawing conclusions based on a particular pattern at one region but not another (see Cunnings & Sturt, 2018 for discussion). By adding 'region' as a fixed effect, we also included a random intercept for trial given that adding 'region' involved two data points from the same trial that are non-independent. A random slop of region was also included by subject, item and trial. The main effects of region and/or group by region interactions will not be discussed below as these effects on their own are not of theoretical interest unless they further interact with distractor or grammaticality.

In a subsequent analysis, we examined whether individual differences in L2 proficiency among L2 speakers and lexical automaticity in both L1 and L2 speakers influenced their reading times or accuracy rates in both dependencies. Each predictor was tested individually across offline and online data. In the L2 proficiency model, proficiency was added as a centered continuous

fixed effect to the maximal model, along with the other experimental fixed effects and their interactions. The same model fitting procedure explained earlier was followed to achieve convergence. Lexical automaticity was treated in the same way, except that the lexical automaticity model included both L1 and L2 speakers and so 'group' was additionally added as a fixed effect together with its interaction with the other experimental factors. All results are reported from the successfully converged models.

3.3 Results

3.3.1 Grammaticality judgements.

3.3.1.1 Group analyses.

3.3.1.1.1 S-V agreement.

Accuracy rates are reported in Table 3 and shown in Figure 1, while summary statistics are reported in Table 4. The results showed a significant main effect of group, as L1 speakers performed more accurately than L2 speakers (estimate = -0.80 (0.151), z = -5.31, $p < .001$). There was also a significant main effect of grammaticality**,** with more correct responses to grammatical than ungrammatical conditions (estimate = -0.91 (0.196), z = -4.67, $p < .001$), and a significant main effect of distractor due to lower accuracy rates for conditions with distractors than conditions without distractors (estimate = 0.83 (0.100), z = 8.24, $p < .001$).

Table 3

Accuracy in Percentages for S-V Agreement in the Grammaticality Judgement Experiment.

	L1 Speakers	L2 Speakers
No Distractor, Grammatical	96 (0.01)	86 (0.01)
No Distractor, Ungrammatical	89 (0.01)	73 (0.01)

Distractor, Grammatical	83 (0.01)	83 (0.01)
Distractor, Ungrammatical	75 (0.01)	74 (0.01)

Note. standard errors in parentheses.

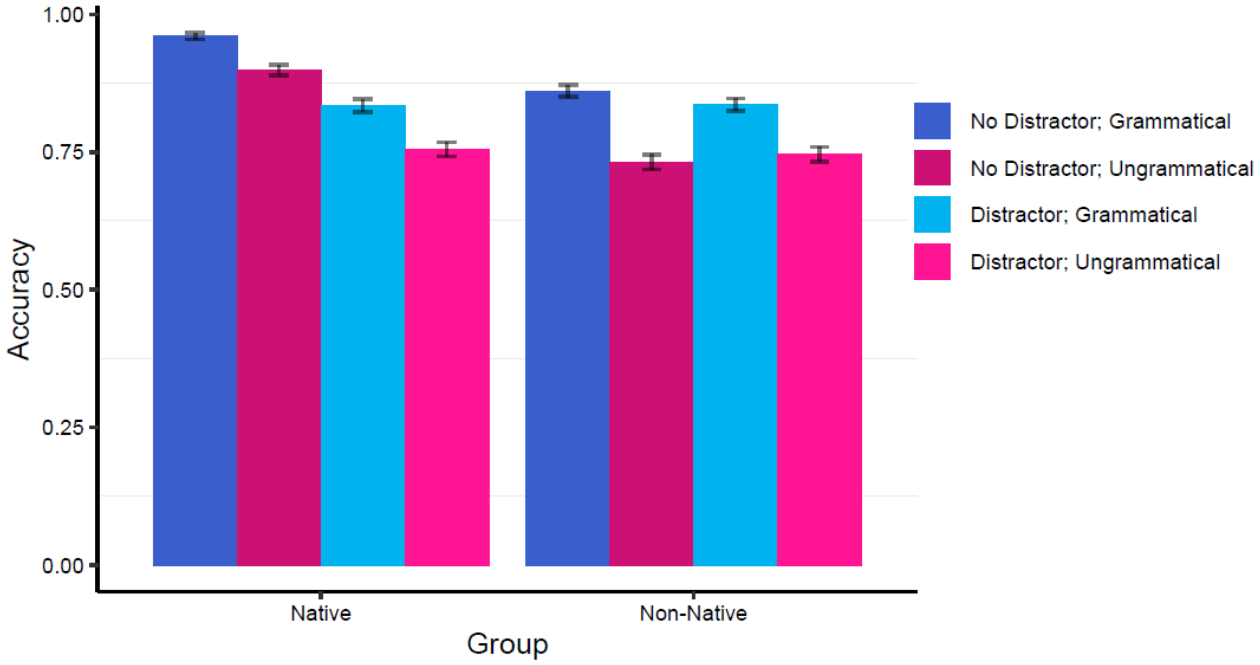


Figure 1. *Accuracy Rates for S-V Agreement Dependencies in the Grammaticality Judgement Experiment in Study 1. Error Bars Represent Standard Errors.*

Table 4

Statistical Analysis for S-V Agreement in the Grammaticality Judgement Experiment.

	Estimates (SE)	*z*	*p*
Group	-0.80 (0.15)	-5.31	**<0.001**
Grammaticality	-0.92 (0.20)	-4.67	**<0.001**
Distractor	0.83 (0.10)	8.25	**<0.001**

Group*Grammaticality	0.11 (0.30)	0.38	0.706
Group*Distractor	-1.58 (0.20)	-8.07	**<0.001**
Grammaticality*Distractor	-0.48 (0.18)	-2.73	**0.006**
Group*Grammaticality*Distractor	0.20 (0.38)	0.53	0.596

Note. significant effects ($p < 0.05$) are in bold.

A significant group by distractor interaction was also found. Nested contrasts showed that distractor effects were significant for the L1 group only, indicating lower accuracy in distractor than no distractor conditions for L1 speakers only (L1 estimate = 1.63 (0.173), z = 9.42, p < .001; L2 estimate = 0.047 (0.110), z = 0.43, p = .66). There was also a significant interaction between grammaticality and distractor (estimate = -0.48 (0.18), z = -2.73, p = .006), but the three-way interaction with group was not significant (estimate = 0.20 (0.38), z = 0.53, p = .596). Nested contrasts, collapsed across distractor/no distractor conditions, indicated higher accuracy in grammatical than ungrammatical conditions, with a larger effect in the no distractor conditions (estimate = -1.15 (0.226), z = -5.12, p < .001) than in the distractor conditions (estimate = -0.664 (0.205), z = -3.23, p = .001).

3.3.1.1.2 Reflexives.

Table 5 and Figure 2 show accuracy rates, while Table 6 summarizes inferential statistics. Analysis revealed a significant main effect of group, with the L1 group being overall more accurate than the L2 speakers (estimate = -0.54 (0.163), z = -3.31, p < .001), and a significant main effect of grammaticality**,** with the grammatical conditions having higher accuracy rates (estimate = -1.14 (0.201), z = -5.67, p < .001). We also observed a significant main effect of distractor, such that conditions without distractors received higher accuracy rates than conditions with distractors (estimate = 0.76 (0.136), z = 5.60, p < .001).

Table 1

Accuracy in Percentages for Reflexives in the Grammaticality Judgement Experiment.

	L1 Speakers	L2 Speakers
No Distractor, Grammatical	97 (0.01)	88 (0.01)
No Distractor, Ungrammatical	81 (0.01)	79 (0.01)
Distractor, Grammatical	87 (0.01)	82 (0.01)
Distractor, Ungrammatical	74 (0.01)	77 (0.01)

Note. standard errors in parentheses.

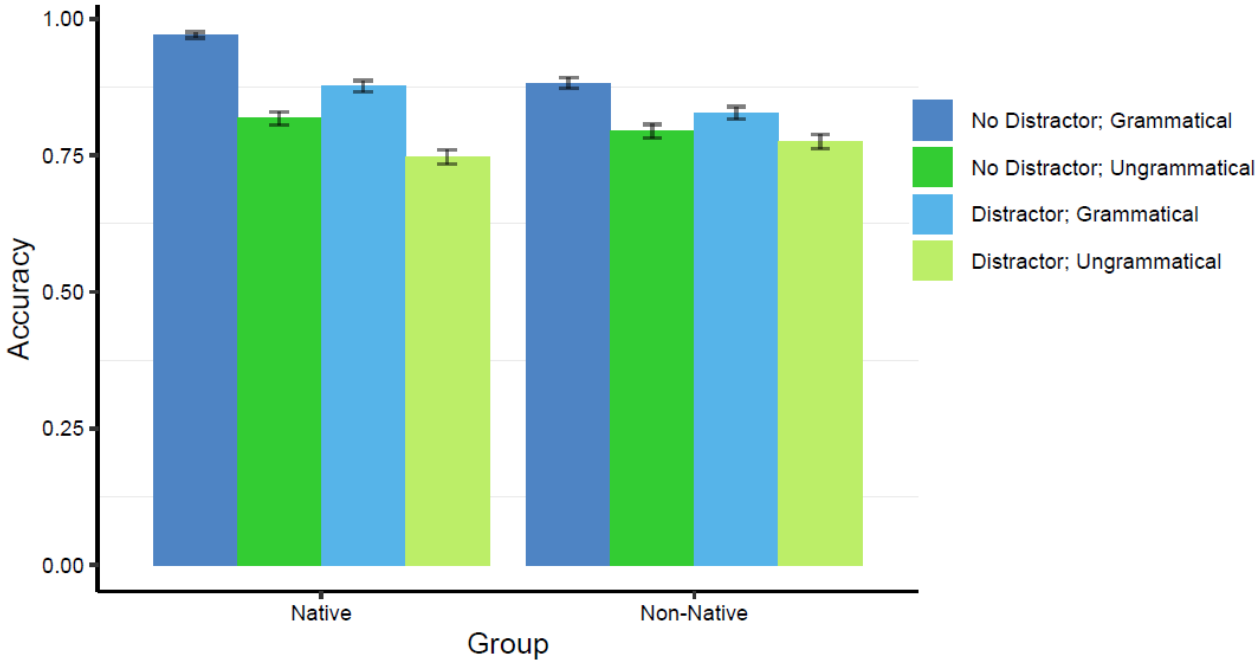


Figure 2. *Accuracy Rates for Reflexive Dependencies in the Grammaticality Judgement Experiment in Study 1. Error Bars Represent Standard Errors.*

Table 2

Statistical Analysis for Reflexives in the Grammaticality Judgement Experiment.

	Estimates (SE)	*z*	*p*
Group	-0.54 (0.16)	-3.32	**0.001**
Grammaticality	-1.14 (0.20)	-5.67	**<0.001**
Distractor	0.77 (0.14)	5.61	**<0.001**
Group*Grammaticality	1.19 (0.28)	4.21	**<0.001**
Group*Distractor	-0.82 (0.17)	-4.72	**<0.001**
Grammaticality*Distractor	-0.75 (0.29)	-2.62	**0.009**
Group*Grammaticality*Distractor	0.82 (0.38)	2.14	**0.033**

Note. significant effects ($p < 0.05$) are in bold.

There was also a significant three-way interaction between group, grammaticality and distractor (estimate = 0.82 (0.38), z = 2.14, p = .033). Nested contrasts suggest significant grammaticality effects, with higher accuracy in grammatical than ungrammatical conditions, for L1 speakers in both distractor/no distractor conditions, with larger effects in the no distractor conditions (no distractor estimate = -2.29 (0.382), z = -6.01, p < .001; distractor estimate = -1.17 (0.269), z = -4.35, p < .001). L2 speakers, however, showed significant grammaticality effects only in the no distractor conditions (no distractor estimate = -0.72 (0.242), z = -3.00, p = .002; distractor estimate = -0.380 (0.303), z = -1.25, p = .210).

3.3.1.2 Individual differences in L2 proficiency.

3.3.1.2.1 S-V agreement.

There was a main effect of proficiency (estimate = 0.07 (0.008), z = 8.08, p < .001) revealing that an increase in proficiency induces an increase in accuracy rates. A significant interaction between proficiency and distractor was also found (estimate = -0.022 (0.010), z = -2.07, p = .037). Figure 3 shows that higher proficiency L2 speakers are more accurate in distractor

than no distractor conditions, while lower proficiency L2 speakers show the opposite trend. Note that, as can be seen in Figure 3, this interaction effect is small, especially when compared to the clear main effect of proficiency.

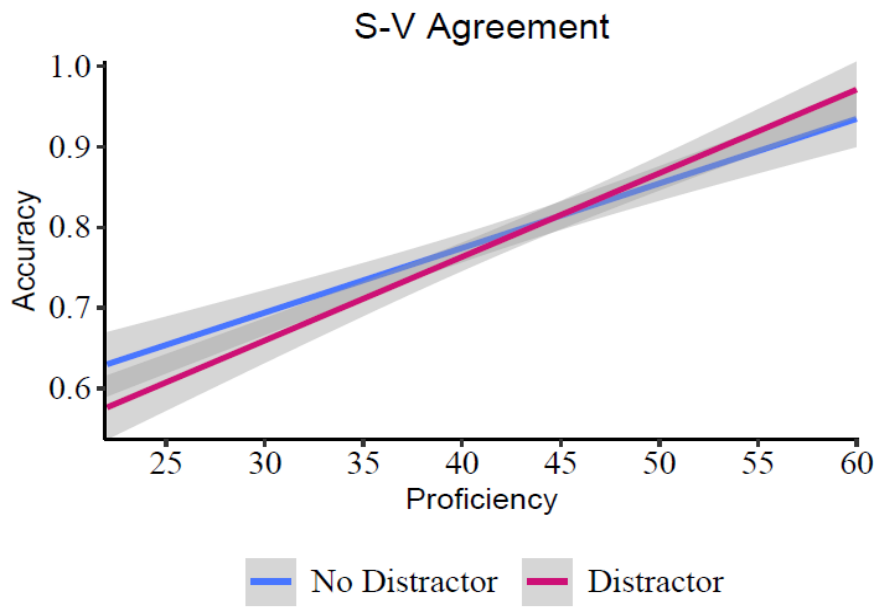


Figure 3. *Effects of Proficiency and Distractor on L2 Speakers' Accuracy rates for S-V Agreement in Study 1.*

3.3.1.2.2 Reflexives.

The main effect of proficiency was significant (estimate = 0.061 (0.007), z = 8.06, p < .001), indicating that accuracy increases as proficiency increases as shown in Figure 4. There was, however, no significant interaction between proficiency and the other predictors (all z < 1.41, all p >.158).

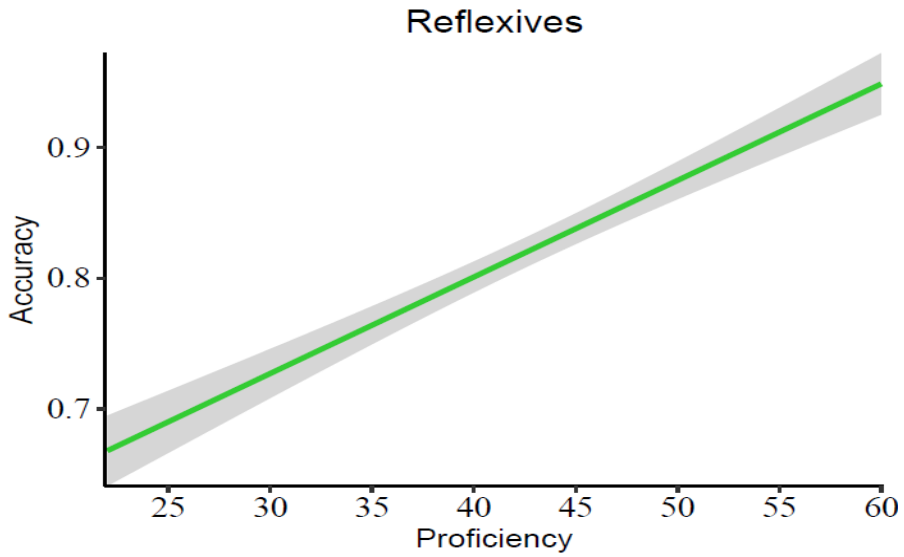


Figure 4. *Effects of Proficiency on L2 Speakers' Accuracy rates for Reflexives in Study 1.*

3.3.1.3 Individual differences in lexical automaticity.

3.3.1.3.1 S-V agreement.

The main effect of lexical automaticity was significant (estimate = -1.15 (0.576), z =-1.99, p = .045) indicating that accuracy rates get lower as lexical processing gets slower. However, there was no significant interaction between lexical automaticity and the other predictors (all z < 1.64, all p >.100).

3.3.1.3.2 Reflexives.

The main effect of lexical automaticity was not significant (estimate = -0.334 (0.506), z = -0.662, p = .508), and there was not also any significant interaction between lexical automaticity and the other predictors (all z < 1.56, all p >.118).

3.3.2 Self-Paced Reading.

3.3.2.1 Group analyses.

3.3.2.1.1 S-V agreement.

Reading times for S-V agreement at the regions of interest are reported in Table 7 while inferential statistics are provided in Table 8.

Table 7

Mean Reading Times (ms) by Condition for S-V Agreement in the Self-Paced Reading Experiment in Study 1.

	Critical Region		Spillover Region	
	L1 Speakers	L2 Speakers	L1 Speakers	L2 Speakers
Grammatical, Matched Distractor	478 (9)	643 (12)	471 (10)	932 (21)
Grammatical, Mismatched Distractor	478 (8)	637 (11)	478 (10)	982 (25)
Ungrammatical, Matched Distractor	512 (10)	693 (19)	550 (13)	1004 (25)
Ungrammatical, Mismatched Distractor	529 (13)	662 (12)	572 (14)	995 (25)

Note. standard errors in parentheses.

Table 8

Statistical Analysis for S-V Agreement in the Self-Paced Reading Experiment in Study 1.

	Estimates (SE)	*t*	*p*
Group	0.45 (0.04)	12.45	<**0.001**
Grammaticality	0.06 (0.01)	8.23	<**0.001**
Distractor	0.01 (0.01)	1.17	0.242
Region	0.16 (0.03)	6.11	<**0.001**
Group*Grammaticality	-0.06 (0.02)	-3.78	<**0.001**
Group*Distractor	-0.01 (0.01)	-0.99	0.323
Grammaticality*Distractor	-0.00 (0.01)	-0.25	0.800
Group*Region	0.30 (0.04)	7.36	<**0.001**
Grammaticality*Region	0.04 (0.01)	3.52	<**0.001**
Distractor*Region	0.01 (0.01)	1.26	0.208
Group*Grammaticality*Distractor	-0.04 (0.03)	-1.20	0.232
Group*Grammaticality*Region	-0.07 (0.03)	-2.64	**0.008**
Group*Distractor*Region	-0.00 (0.02)	-0.02	0.988

Grammaticality*Distractor*Region	-0.00 (0.02)	-0.15	0.882
Group*Grammaticality*Distractor* Region	-0.06 (0.05)	-1.35	0.177

Note. significant effects ($p < 0.05$) are in bold.

Reading times revealed a significant main effect of group, as L2 speakers generally had longer reading times. There was also a significant main effect of grammaticality, driven by longer reading times in the ungrammatical sentences relative to the grammatical sentences. This ungrammaticality effect was not attenuated by matching distractors as expected, especially for L2 speakers whose readings times did not largely diverge between ungrammatical conditions as shown in Figure 5, and neither the main effect of distractor, nor any interactions with distractor, were significant.

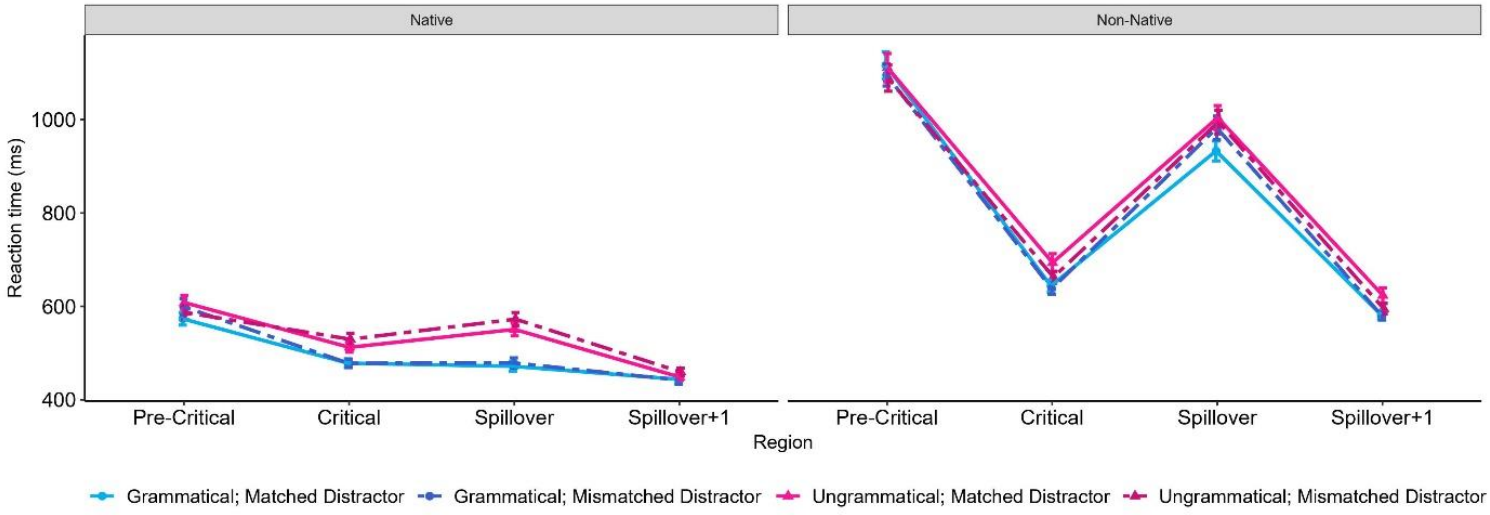


Figure 5. *Reading Times for S-V Agreement Dependencies in the Self-Paced Reading Experiment in Study 1. Error Bars Represent Standard Errors.*

We also found a significant three-way interaction between group, grammaticality and region**.** Nested contrasts revealed that grammaticality effects were significant only for the L1 speakers at the critical region (L1 estimate = 0.049 (0.013), t = 3.80, p < .001; L2 estimate = 0.027 (0.013), t = 1.93, p = .064), but they were significant for both groups at the spillover region,

though the L1 estimate was numerically larger (L1 estimate = 0.126 (0.013), t = 9.37, p < .001; L2 estimate = 0.035 (0.015), t = 2.32, p = .029).

3.3.2.1.2 Reflexives.

A summary of mean reading times at the critical and spillover regions and the inferential statistics are provided in Table 9 and Table 10, respectively.

Table 9

Mean Reading Times (ms) by Condition for Reflexives in the Self-Paced Reading Experiment in Study 1.

	Critical Region		Spillover Region	
	L1 Speakers	L2 Speakers	L1 Speakers	L2 Speakers
Grammatical, Matched Distractor	528 (12)	828 (18)	463 (9)	652 (11)
Grammatical, Mismatched Distractor	521 (11)	850 (18)	478 (9)	656 (12)
Ungrammatical, Matched Distractor	612 (20)	1037 (31)	569 (15)	875 (22)
Ungrammatical, Mismatched Distractor	670 (23)	1133 (36)	598 (15)	901 (23)

Note. standard errors in parentheses.

Table 10

Statistical Analysis for Reflexives in the Self-Paced Reading Experiment in Study 1.

	Estimate (SE)	*t*	*p*
Group	0.43 (0.04)	11.9	**<0.001**
Grammaticality	0.14 (0.01)	10.8	**<0.001**
Distractor	0.02 (0.01)	3.46	**0.001**
Region	-0.11 (0.03)	-4.26	**<0.001**
Group*Grammaticality	0.06 (0.02)	2.89	**0.004**
Group*Distractor	-0.01 (0.01)	-0.38	0.707
Grammaticality*Distractor	0.02 (0.02)	1.48	0.139
Group*Region	-0.12 (0.03)	-4.05	**<0.001**
Grammaticality*Region	0.07 (0.01)	5.18	**<0.001**

Distractor*Region	0.00 (0.01)	0.26	0.792
Group*Grammaticality*Distractor	-0.01 (0.03)	-0.33	0.744
Group*Grammaticality*Region	0.02 (0.03)	0.68	0.495
Group*Distractor*Region	-0.04 (0.03)	-1.46	0.145
Grammaticality*Distractor*Region	-0.01 (0.03)	-0.30	0.764
Group*Grammaticality*Distractor*Region	0.02 (0.06)	0.30	0.767

Note. significant effects ($p < 0.05$) are in bold.

For reading times, there was a main effect of group, as L2 speakers were generally slower readers than L1 speakers. There were also significant main effects of grammaticality and distractor, with longer reading times in ungrammatical than grammatical sentences and in sentences with mismatching distractors than with matching distractors. As shown in Figure 6, the distractor effect is largely restricted to ungrammatical sentences, and is suggestive of facilitatory interference in ungrammatical sentences only. However, the interaction between grammaticality and distractor did not reach statistical significance.

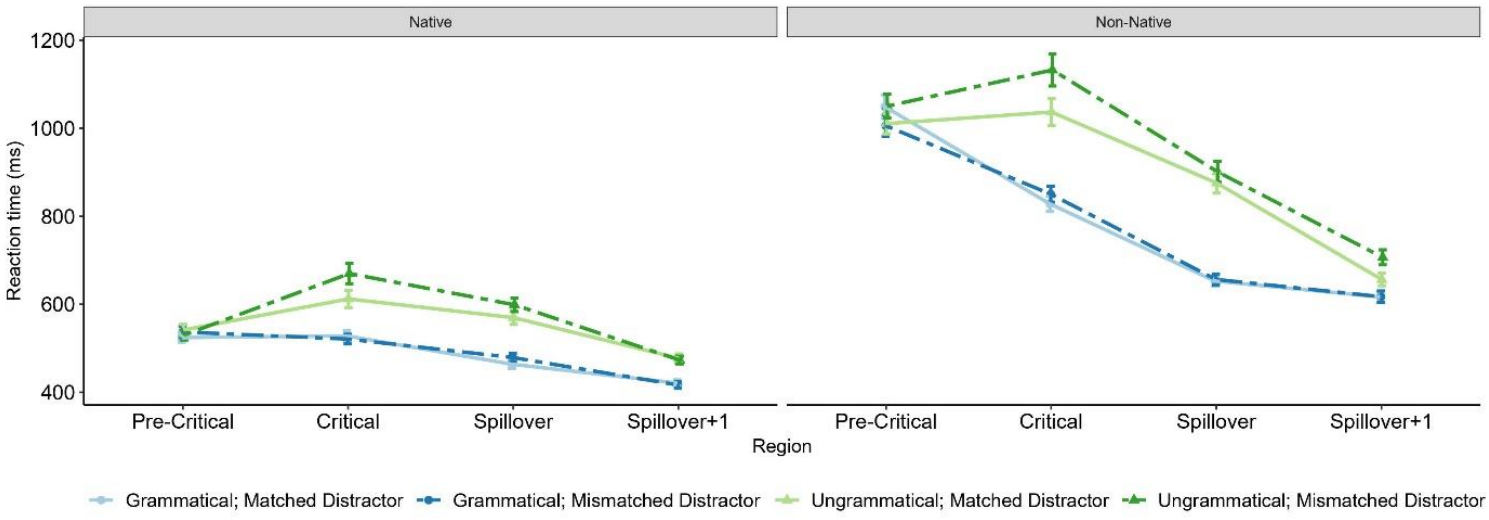


Figure 6. *Reading Times for Reflexive Dependencies in the Self-Paced Reading Experiment in Study 1. Error Bars Represent Standard Errors.*

The group by grammaticality interaction was significant. Nested contrasts indicated significant grammaticality effects for both groups, with larger effects for the L2 speakers (L1 estimate = 0.111 (0.014), t = 7.70, p < .001; L2 estimate = 0.168 (0.018), t = 9.37, p < .001). A significant grammaticality by region interaction was also found. Nested contrasts showed that grammaticality effects were significant at both the critical and spillover regions, with larger effects at the spillover region (critical region estimate = 0.103 (0.015), t = 6.77, p < .001; spillover region estimate = 0.176 (0.013), t = 12.99, p < .001).

3.3.2.2 Individual differences in L2 proficiency.

3.3.2.2.1 S-V agreement

There was a significant main effect of proficiency (estimate = -0.0006 (0.0002), t = -2.11, p = .035), showing that reading times get shorter as proficiency increases. There was also a significant interaction between grammaticality and proficiency (estimate < 0.001 (0.0001), t = 2.91, p = .007). Figure 7 shows that higher proficiency L2 speakers had longer reading times in ungrammatical conditions compared to grammatical conditions while lower proficiency L2 speakers behaved differently. No other significant interactions of theoretical interest with proficiency were additionally observed (all t < .85, all p > .392).

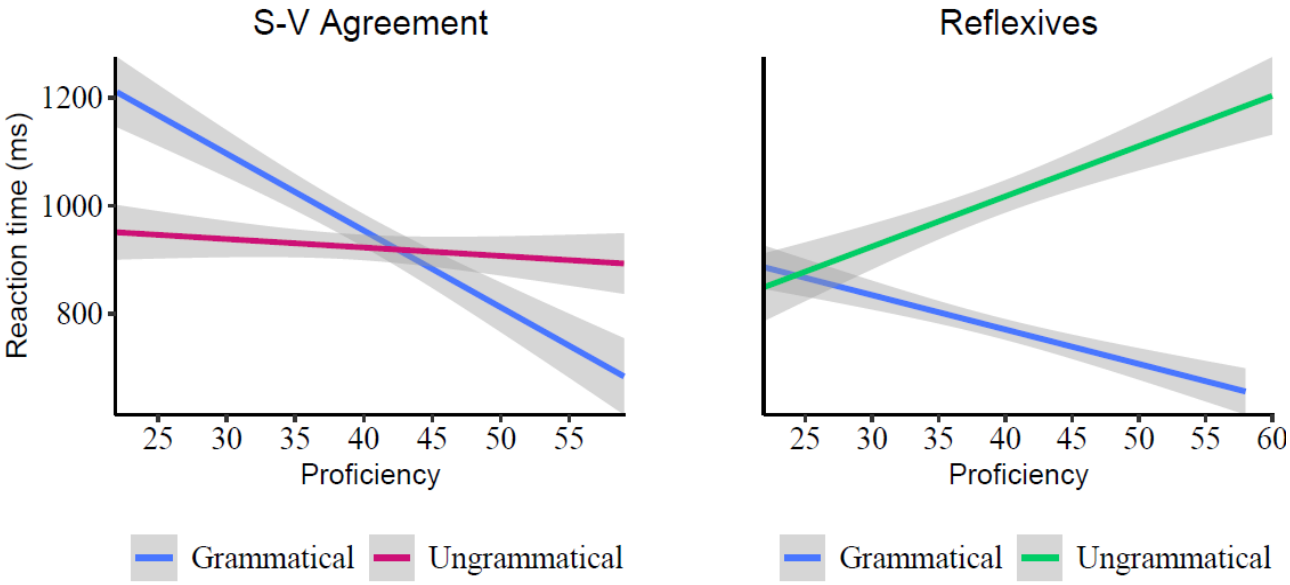


Figure 7. *Effects of Proficiency and Grammaticality on L2 Speakers' Reading Times in Study 1.*

3.3.2.2.2 Reflexives.

The main effect of proficiency was not significant (estimate < 0.001 (0.002), t = 0.127, p = .899). However, there was a significant interaction between grammaticality and proficiency (estimate = 0.006 (0.001), t = 4.59, p < .001). As illustrated in Figure 7, the effect of grammaticality gets larger as proficiency increases. We did not find any other significant interactions of theoretical interest with proficiency (all t < 1.07, all p > .285).

3.3.2.3 Individual differences in lexical automaticity.

3.3.2.3.1 S-V agreement.

No significant main effect of lexical automaticity was found (estimate = -0.012 (0.127), t = - 0.095, p = .92) nor any significant interaction between lexical automaticity and any of the other predictors (all t < 1.84, all p > .068).

3.3.2.3.2 Reflexives.

The main effect of lexical automaticity was not significant (estimate = -0.07 (0.134), t = -0.57, p = .56). There was, however, a significant five-way interaction between region, lexical automaticity, grammaticality, distractor, and group (estimate = 1.06 (0.469), t = 2.26, p = .023).

As illustrated in Figure 8, L2 speakers with low levels of lexical automaticity (i.e., with high CV values that are far greater than .1 on the lexical automaticity scale) tend to show larger grammaticality effects at the critical region as compared to L2 speakers with high levels of lexical automaticity (i.e., with low CV values that are closer to .1 on the scale). They might also be slightly less influenced by the matching distractors relative to the highly automatised L2 speakers. On the other hand, L1 speakers with low levels of lexical automaticity appear to be less sensitive to grammaticality effects relative to the highly automatised participants. They also seem to be more influenced by the matching distractors across grammatical and ungrammatical conditions. At the spillover region, all L2 speakers showed generally similar grammaticality effects, but those less automatized tend to be more influenced by the matching distractors in ungrammatical conditions. L1 speakers with high levels of automaticity appear to show clearer grammaticality effects and might be marginally affected by the matching distractors in ungrammatical conditions compared to less automatised participants.

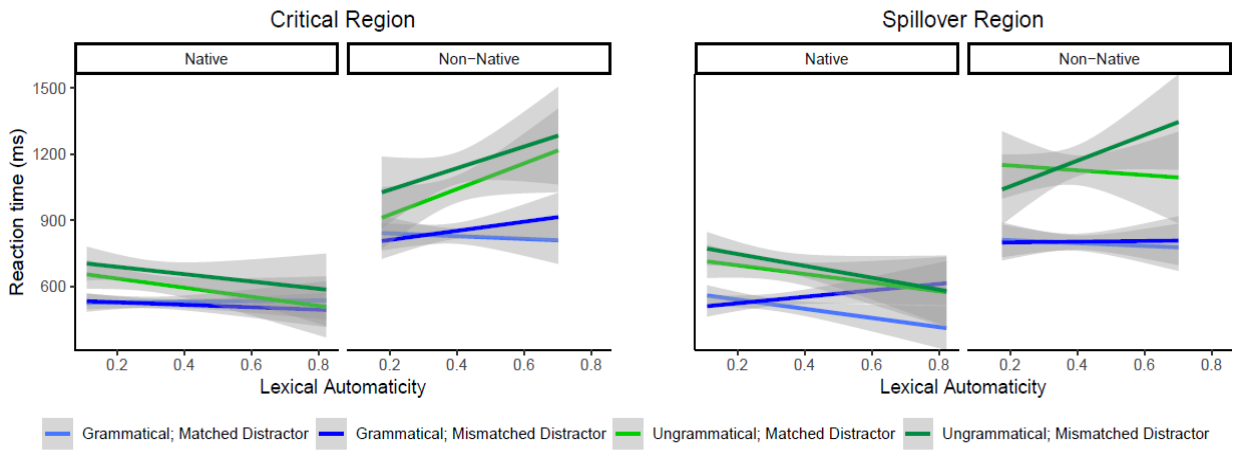


Figure 8. *Interaction Effect Between Grammaticality, Distractor, Group, Lexical Automaticity and Region* on L2 Speakers' Reading Times *for Reflexive Dependencies in Study 1.*

3.4 Discussion

3.4.1 Grammaticality judgements.

The results of the grammaticality judgements showed generally high accuracy of L1 and L2 speakers in both S-V agreement and reflexives, suggesting both groups correctly identified the permissible and impermissible conditions according to structural constraints. It is clear, therefore, that L2 speakers have adequate grammatical knowledge of these dependencies. On closer inspection, both groups' judgements also seem biased towards 'grammatical' responses, as evidenced by significant main effects of grammaticality, perhaps suggesting response bias that is sometimes observed in binary forced choice tasks (see Felser et al., 2009; Hammerly et al., 2019). This can explain the difference in accuracy rates in both groups between grammatical and ungrammatical sentences. This difference or response bias also seemed larger in the 'no distractor' conditions.

Conditions containing distractors, whether grammatical or ungrammatical, also had lower accuracy than conditions without, especially in L1 speakers, which may just indicate lower acceptance rates for longer sentences. Finding unexpectedly L1 speakers more influenced by distractors than L2 speakers stands against what some L2 models might predict (Clahsen & Felser, 2006a, 2018; Cunnings, 2017). It is difficult however to determine whether similarity-based interference may have played any role here, as we did not manipulate the distractor's properties in this offline experiment, and so we do not have conditions with matching and mismatching distractors to see how judgements may differ accordingly and compared to the baseline conditions.

Finally, though overall the L2 group was slightly less accurate than L1 speakers, L2 speakers' performance on the task was also influenced by individual differences, particularly in L2 proficiency. In both dependencies, accuracy rates increased as proficiency increased.

Increasing lexical automaticity also contributed to an increase in accuracy rates for S-V agreement.

Online data would provide further evidence of whether participants abide by structural constraints, as shown overall in the offline task, and whether they are affected by similarity-based interference effects or individual differences.

3.4.2 Self-paced reading.

The results showed that L2 speakers' performance largely patterned with L1 speakers' in that their reading times were guided by structural constraints, suggesting that both groups apply a qualitatively similar parsing mechanism. Specifically, they showed significant grammaticality effects on both dependencies, but the size of effects differed across the two dependencies. While L2 speakers' grammaticality effects were relatively smaller for S-V agreement compared to L1 speakers, grammaticality effects were bigger for L2 speakers for reflexives relative to L1 speakers. The results of proficiency analyses may explain this relative difference. That is, while almost all L2 speakers showed reliable grammaticality effects for reflexives across different proficiency levels with larger effects as proficiency increased as shown in Figure 7, only higher proficiency L2 speakers demonstrated grammaticality effects in the target direction for S-V agreement, with longer reading times in ungrammatical than grammatical conditions.[16]

Effects of distractors could only be observed in reflexives for both groups similarly, such that conditions with matching distractors were read faster than those with mismatching

[16] To assess further proficiency effects, L2 participants were binned into 2 proficiency groups, and additional analyses were carried out separately for those who fell within the 'advanced' proficiency category (proficiency score=>40, n=108) and those who scored lower (n=68). As these two groups stand at the upper and lower ends of the proficiency spectrum, differences in development are most likely to be noticed. The results were the same as in the main analysis. However, only lower proficiency L2 speakers showed statistically smaller grammaticality effects compared to L1 speakers, while advanced L2 speakers and L1 speakers did not differ significantly, suggesting developmental changes in L2 morphosyntactic agreement processing as a function of proficiency.

distractors. From Figure 6, it can be seen that this effect was largely driven by ungrammatical conditions as there were nearly no differences between grammatical conditions. This descriptive finding seems to be an indication of similarity-based interference, namely facilitatory interference. The lack of distractor effects for S-V agreement was surprising, and we return to this in the General Discussion. Importantly for our present purposes however, distractor effects did not interact with group, as such we found no evidence to suggest either shallow L2 processing or increased susceptibility to interference in L2 speakers.

The degree of lexical automaticity also appeared to have no clear effects on S-V agreement processing. We also did not find consistent effects of lexical automaticity on reflexive processing, but only a complicated five-way interaction of group, grammaticality, distractor and region with lexical automaticity that is difficult to interpret. The only consistent pattern observed across both the critical and spillover regions as a result of this complex interaction was that L1 speakers with higher lexical automaticity had larger grammaticality effects than less automatized L1 speakers, as demonstrated in Figure 9. Considering the main analysis that found a significant interaction between group and grammaticality with L1 speakers being less sensitive to grammaticality effects relative to L2 speakers, the size of grammaticality effects displayed by L1 speakers might have something to do with accuracy in lexical recognition and efficiency of lexical processing. Other interaction effects, however, were not consistent and hence it is difficult to draw any further conclusions. Note also that the shaded areas for all the conditions at the lexical automaticity scale in Figure 9 demonstrate wide confidence intervals, suggesting instability and a high degree of uncertainty.

Finally, the fact that L2 speakers showed smaller grammaticality effects compared to L1 speakers during S-V agreement processing relative to reflexive dependencies may suggest reduced L2 sensitivity to number than gender. However, this reduced sensitivity to number might be related either to L2 difficulty in representing number itself, or alternatively it might be related

to difficulty in computing agreement. To assess whether this effect relates to the representation of number itself we ran Study 2, which examined reflexives but using a number, rather than gender, manipulation.

CHAPTER FOUR

4. Study 2

4.1 Introduction

Study 1 revealed L2 speakers' reduced sensitivity to grammaticality effects in S-V agreement relative to L1 speakers, but this L1/L2 difference was not found for reflexives, which used gender manipulation, suggesting potentially less sensitivity to number than gender.

In L2 research, number and gender features have often been investigated in the domain of agreement, with most research have been devoted to addressing different theoretical issues in L2 acquisition and development. Several behavioural studies showed that gender agreement in particular, such as between determiners and nouns or adjectives and nouns, poses a significant challenge to L2 speakers, especially because it is more subject to cross-linguistic variation (e.g., Franceschina, 2001, 2005; Hawkins & Franceschina, 2004; McCarthy, 2008; White, Valenzuela, Kozlowska-Macgregor, & Leung, 2004). This difficulty with gender agreement has often been addressed in terms of how L2 grammatical development is influenced by the properties of L2 speakers' native language and how the (un)availability of uninterpretable parameterized features, such as grammatical gender, in the L1 may determine ultimate attainment in the L2 following models such as the Failed Functional Feature Hypothesis (Hawkins & Chan, 1997) and, more recently, the Interpretability Hypothesis (Tsimpli & Dimitrakopoulou, 2007), among others. Franceschina (2005), for instance, found that advanced L2 speakers of Spanish from L1 languages with grammatical gender (e.g., Arabic, French, Italian, etc.) peformed better on gender agreement in several tasks than advanced English L2 speakers of Spanish whose L1 lacks grammatical gender. English L2 speakers also generally performed better with number agreement than gender agreement. Consequently, ultimate L2 attainment was argued to be affected by L1 syntactic features inventory. Unlike Franceschina, White et al. (2004) found no effects for L1 properties,

but did find a significant effect for proficiency. Specifically, the lower proficiency L1 French and L1 English speakers of L2 Spanish were the least accurate regardless of whether the L1 is grammatically gendered or not. Additionally, the lower proficiency L2 participants, including L1 French speakers, were also the only group to show an advantage for number over gender since the gender of lexical items also varies in French and Spanish. However, multiple measures showed that advanced L2 speakers, both French and English native speakers, were indistinguishable from Spanish native speakers. This suggests that agreement can be processed natively in the L2 at advanced proficiency without any effects from the L1. More recently, a number of ERP studies also revealed that processing gender and number agreement is qualitatively similar, as suggested by the P600 effects for both features for the advanced L2 speakers, with proficiency playing a significant role in achieving a native-like grammaticality sensitivity (e.g., Alemán Bañón & Rothman, 2016; Alemán Bañón, Fiorentino, & Gabriele, 2018; Dowens, Vergara, Barber, & Carreiras, 2010; Gabriele, Fiorentino, & Alemán Bañón, 2013; cf. Foucart & Frenck-Mestre, 2011). In general, while earlier studies observed inconsistent findings regarding whether gender and number features are represented or processed similarly, recent research showed similar processing profiles for the two features.

After briefly discussing how number and gender are usually considered and investigated in L2 research, it is necessary to clarify here that the discussion above pertains to 'grammatical gender', which is a formal property distinct from 'semantic gender', which we manipulated between a reflexive and its antecedent in Study 1. That is, semantic gender is typically used for words referring to animate entities based on the semantic notion of biological gender, and hence it reflects a transparent relationship between the biological sex of the referents and the gender of nouns and pronouns, such as reflexives, whereas grammatical gender is arbitrarily assigned to words whether they have biological sex or not (Barber & Carreiras, 2005; Ritter, 1993; White et al., 2004). From a theoretical syntactic perspective, languages with grammatical gender, such as

Spanish, have interpretable gender features on nouns[17] and uninterpretable gender features on determiners and adjectives that should be checked or valued via agreement, in the sense of Chomsky's (1995) feature checking, with the interpretable gender features of nouns (Carstens, 2000; Franceschina, 2001; Montrul, Foote, & Perpiñán, 2008; White et al., 2004). It is this syntactic mechanism for gender agreement that has been argued to cause difficulty, especially if gender agreement (formal uninterpretable gender feature) is not represented in L2 speakers' L1 grammars at an abstract syntactic level (Hawkins & Franceschina, 2004; Montrul et al., 2008). However, semantic gender agreement between anaphoric entities and their antecedents do not follow the same syntactic agreement mechanism (Osterhout & Mobley, 1995). Therefore, drawing conclusions from the literature discussed above about the findings of Study 1, where L2 speakers appeared to perform better on gender than number, is difficult due to the different types of gender and syntactic categories of agreeing elements. Study 1 findings, in which semantic gender was manipulated, cannot thus be construed as contradicting previous research on grammatical gender.

As English lacks grammatical gender, we cannot also examine gender agreement in the verb phrase or in the noun phrase to determine if the reduced grammaticality effects in S-V agreement are attributed to the type of feature manipulated or to agreement computation. Thus, we instead manipulated number in reflexive dependencies in Study 2 to investigate whether Study 1 results indicate reduced sensitivity to number across the board, or whether it is something particularly related to S-V agreement dependency processing. If L2 speakers have difficulty representing number in general, they should show reduced grammaticality effects irrespective of

[17] The notion of grammatical gender on nouns being interpretable in the sense that they contribute to semantic interpretation may seem somewhat counterintuitive since gender distinction has no conceptual basis and is arbitrary. It also does not affect interpretation in any way. However, it is a conceptual necessity in the Minimalist Program (Chomsky, 1995) that features, particularly phi (agreement) φ-features, are generalised as pairs so that uninterpretable features are matched with their equivalent interpretable features for the process Agree to operate and for the uninterpretable features thereafter to be checked and valued by their equivalent interpretable features. A discussion of this syntactic analysis and grammatical gender is beyond the scope of the current study, however. See Carstens (2000) for relevant remarks.

dependency. If however this L2 difficulty is restricted to agreement, L2 speakers may not show reduced grammaticality effects for number when tested using reflexives.

The study additionally aimed to examine whether number manipulation in reflexive dependencies would trigger any interference effects. Since the previous few studies (Felser & Cunnings, 2012) only looked at the grammaticality effects for reflexives with gender manipulation, Study 2 thus contributes to the existing literature as no known research has investigated number congruency in reflexives for L2 speakers.

Similar to Study 1, participants completed an offline grammaticality judgement task along with the self-paced reading experiment. Individual differences in proficiency were also examined, but lexical automaticity was not included as it did not have much effect in Study 1.

4.2 Method

4.2.1 Participants.

168 Saudi Arabic L2 English speakers and 171 L1 English speakers were recruited in total. L1 English speakers were approached through media advertising and the University of Reading student community, and they participated voluntarily or received course credit for their participation. All Arabic L2 English speakers who voluntarily participated in this study had also participated in the first study. The interval between their first and second participation was between six and eleven months.

Using the same inclusion criteria as in the first study, we excluded four Arabic L2 English speakers and five L1 English speakers because they scored less than 75% correct on the self-paced reading comprehension questions which may indicate diminished attention. Two additional L1 English speakers were excluded from the study due to their fast median reaction times during

the grammaticality judgement experiment, which fell below the threshold value of 1500 ms that was set as a measure of their attention.

This ultimately led to the inclusion of 164 Saudi-Arabic L2 English speakers (29 males, mean age = 31, range = 19-44) and 164 L1 monolingual English speakers (21 males, mean age = 25, range = 18-60); 149 of these were British, 17 were American, four were Canadian and one was Australian. For background information, including the age of onset of learning English and L2 proficiency, among other relevant details, please refer to Study 1 for an overview of the L2ers' group.

4.2.2 Materials.

4.2.2.1 Grammaticality Judgement Task (GJT).

24 sets of experimental sentences were created as in (38). Grammaticality (grammatical vs. ungrammatical) and the presence of a distractor (distractor vs. no distractor) were manipulated as two experimental factors. The reflexive was kept plural across all conditions, so grammaticality was manipulated by varying the number feature of its antecedent, the subject. As singular reflexives are gendered and conflate two features, gender and number, a plural rather than a gendered singular reflexive was used in all conditions to avoid the possibility that the gender feature in the latter may come into play and have an effect. In conditions with distractors, the distractor was embedded in a prepositional phrase, mismatching the reflexive number in grammatical conditions but matching it in ungrammatical conditions, as in the previous offline experiment. The 24 experimental items were distributed across four lists in a Latin-square design and combined with 60 filler items, half of which were grammatical. The full set of stimuli is provided in Appendix D.

(38) Reflexives

a. *No Distractor, Grammatical*

The boys hurt themselves yesterday afternoon.

b. *No Distractor, Ungrammatical*

*The boy hurt themselves yesterday afternoon.

c. *Distractor, Grammatical*

The boys near the girl hurt themselves yesterday afternoon.

d. *Distractor, Ungrammatical*

*The boy near the girls hurt themselves yesterday afternoon.

4.2.2.2 Self-Pased Reading Task (SPR).

To examine interference effects and the magnitude of grammaticality effects in comparison to previous experiments, this study employed number manipulation in the comprehension of reflexive dependencies. Grammaticality (grammatical vs. ungrammatical) and distractor match (distractor-match vs. distractor mismatch) were manipulated in 24 maximally similar configurations to those used in the previous online experiment as shown in (39). The reflexive was always plural 'themselves', so reading times at the critical region across conditions will not be affected by different word lengths. Grammaticality depended on number agreement between the reflexive and its antecedent, requiring a plural subject in grammatical conditions and a singular subject in ungrammatical ones. Interference effects were tested by manipulating the distractor number, so that it either matched or mismatched the reflexive across grammatical and ungrammatical conditions.

In addition to the 24 item sets, 12 experimental items of two baseline conditions were also added. In these two conditions, the reflexive was also kept plural and there was no intervening distractor between the reflexive and the subject as shown in (40). The reason for including them was to assess the effects of grammaticality without distractors, particularly because some readers might interpret the reflexive 'themselves' as a non-gendered singular reflexive (Foertsch &

Gernsbacher, 1997; Speyer & Schleef, 2019). This possibility could impact their reading times in the ungrammatical conditions (39c/d), as it may sound acceptable to retrieve the singular subject as the antecedent of ‘themselves’. Accordingly, we may not be able to determine if their reading times, for example in (39c), are influenced by this issue or by interference effects. The baseline conditions (40a/b) can rule out this ambiguity as they would demonstrate how readers perceive the reflexive ‘themselves’ with the singular and plural subject, such that finding grammaticality effects here would suggest ’themselves’ is indeed treated as plural.

(39) Reflexives

a. *Grammatical, Distractor Match*

The girls near the boys accidently hurt themselves while playing in the garden.

b. *Grammatical, Distractor Mismatch*

The girls near the boy accidently hurt themselves while playing in the garden.

c. *Ungrammatical, Distractor Match*

*The girl near the boys accidently hurt themselves while playing in the garden.

d. *Ungrammatical, Distractor Mismatch*

*The girl near the boy accidently hurth themselves while playing in the garden.

(40)

a. *Grammatical*

The boys carefully washed themselves using soap before having dinner.

b. *Ungrammatical*

*The boy carefully washed themselves using soap before having dinner.

The 36 experimental items were mixed with 90 filler items in a pseudo-randomised Latin-square design, such that each participant read a total of 126 sentences. The fillers comprised grammatical sentences that mimicked the target sentences but began with singular subjects, to prevent participants from developing a strategic behavior whereby they associate singular subjects with ungrammatical sentences. Additionally, some fillers resembled the experimental sentences but lacked reflexives, featuring instead main verbs without agreement morphology or forms of the verb *be*. The fillers also included items with singular gendered reflexives and different types of anaphors in different positions, as well as sentences with relative clauses and other structurally distinct types of sentences. After each sentence, participants answered a yes/no comprehension question that had been counterbalanced across conditions, to ensure that they had read the sentence for meaning. The entire set of stimuli can be found in Appendix E.

4.2.3 Procedure.

The study was piloted with four native English speakers before data collection to validate the materials and test randomisation. The procedure was the same as in Study 1. The experiments were initially administered using the web-based IbexFarm software (Drummond, 2013). However, due to the shutdown of IbexFarm in the middle of the data collection process, we had to switch to PCIbex (Zehr & Schwarz, 2018) as a compatible platform to complete data collection[18]. This study was also completed in two sessions with at least one week between sessions. The self-paced reading experiment was conducted in the first session and lasted around 20-30 minutes for L1 speakers and 30-40 minutes for L2 speakers. The second session included the grammaticality judgement experiment and lasted approximately 15-20 minutes for L1 speakers and 20-25 for L2 speakers. As all L2 speakers completed the proficiency test in the first

[18] All L2 speakers completed the tasks using IbexFarm. 40 L1 speakers completed the self-paced reading and 21 L1 speakers completed the grammaticality judgement task in IbexFarm, but the remaining L1 data were collected using PCIbex. Beyond the server on which the experiments were hosted, the two versions of the experiment were identical.

study, they were not asked to complete it again. We also did not include the lexical automaticity measure since we did not find any clear effects in Study 1.

4.2.4 Data analysis.

As previously noted, data were analysed only from participants who achieved at least 75% accuracy on the self-paced reading comprehension questions and had a median reaction time of over 1500 ms on the grammaticality judgement task. These criteria were used as a proxy to ensure that participants paid sufficient attention in both experiments.

Data analysis was performed in the R software environment (version 4.0.3) (R Development Core Team, 2020). A binomial generalized mixed-effects model with the lme4 package (Baayen et al., 2008; Bates et al., 2015) was used to analyse grammaticality judgement accuracy data for binary judgement responses. Fixed effects of group (L1/L2), grammaticality (grammatical/ungrammatical) and distractor's presence (distractor/no distractor) were sum coded (-0.5/0.5) and included in the analysis along with their interactions in a maximal random effects model with by-subject and by-item random intercepts and random slopes for all the fixed effects. We began our analysis by fitting a model with a maximal random effects structure with by-subject and by-item random intercepts, as well as random slopes for each fixed effect (Barr et al., 2013). If the maximal model failed to converge, we refitted the model without correlation parameters. If this model was still unable to converge, the least variance-causing random effects were incrementally removed until convergence was achieved. In case of finding an interaction between variables, a follow-up analysis was conducted using nested contrasts of the explanatory variable separately at levels of the other interacting variable(s), which were then included along with the other effects in a maximal random mixed effects model following the same model fitting process as in the main analysis until successful convergence.

For the self-paced reading experiment, we conducted separate analyses for each experimental set. We had two sets of experimental items: one set consisting of 12 items that merely compared baseline grammatical and ungrammatical conditions with plural reflexives to assess sensitivity to

grammaticality, and another set consisting of 24 items that examined grammaticality and interference effects with matching and mismatching distractors. Reading times shorter than 100 ms or longer than 10,000 ms were first discarded as they may reflect distractions or inattention. This accounted for less than 0.01% of the data for both groups. Reading times from both sets were then log-transformed to eliminate skewness, and analysed using linear mixed effects models with the lme4 package (Baayen et al., 2008; Bates et al., 2015). Our analysis of the two experimental sets examined reading times at the reflexive 'themselves' (the critical region) and the next word (the spillover region). For the first experimental set, fixed effects included sum coded (-0.5/0.5) effects of group (L1/L2), grammaticality (grammatical/ungrammatical) and region (critical/spillover). The same fixed effects were used in the analysis of the second experimental set, with the addition of distractor matching effect (match/mismatch) as another fixed effect. As the inclusion of 'region' as a fixed effect involved two non-independent data points from a single trial, the maximal random effects model included a random intercept for trial and a by-trial random slope for region (see Cunnings & Sturt, 2018 for discussion). The same iterative procedure used for fitting a model for accuracy data was also followed for reading times until convergence was reached. Like in Study 1, the main effect of region and/or the group by region interaction were not discussed unless they interacted with a variable of interest, such as grammaticality and/or distractor, as they are not of theoretical significance on their own.

A subsequent analysis of L2 proficiency was also conducted to examine whether differences in proficiency had an impact on the resulting L2 processing pattern and L2 speakers' grammaticality judgements. The Oxford Quick Placement Test's proficiency scores of all L2 speakers in this study[19], who had previously completed the test in Study 1, were included as a centered continuous fixed effect along with the other fixed effects and their interactions in a mixed-effects model. As no measure of lexical automaticity was included in this study, primarily due to the absence of clear effects observed in Study 1, lexical automaticity scores for L1 speakers are unavailable. Therefore, no further analysis

[19] The participants' proficiency scores ranged from 22 to 60.

was conducted for lexical automaticity. The following results are reported from models that successfully converged.

4.3 Results

4.3.1 Grammaticality judgements.

4.3.1.1 Group analyses.

Accuracy rates and a summary of the statistical analysis are provided in Tables 11 and 12, respectively.

Table 11

Accuracy in Percentages for Reflexive in the Grammaticality Judgement Experiment in Study 2.

	L1 Speakers	L2 Speakers
Grammatical, Distractor	87 (0.01)	88 (0.01)
Grammatical, No Distractor	97 (0.01)	91 (0.01)
Ungrammatical, Distractor	43 (0.01)	67 (0.01)
Ungrammatical, No Distractor	46 (0.01)	71 (0.01)

Note. standard errors in parentheses.

Table 12

Statistical Analysis for Reflexives in the Grammaticality Judgement Experiment in Study 2.

	Estimates (SE)	*t*	*p*
Group	1.53 (0.27)	2.42	**0.015**
Grammaticality	0.05 (0.01)	-11.38	**<0.001**
Distractor	2.11 (0.22)	7.32	**<0.001**
Group*Grammaticality	11.75 (4.03)	7.19	**<0.001**
Group*Distractor	0.51 (0.11)	-3.16	**0.002**

Grammaticality*Distractor	0.36 (0.11)	-3.43	**0.001**
Group*Grammaticality*Distractor	4.12 (1.64)	3.56	**<0.001**

Note. significant effects ($p < 0.05$) are in bold.

As can be seen, there was a significant main effect of group (estimate = 0.427 (0.176), z = 2.42, p = .015), with the L2 group having higher accuracy rates than the L1 group. Significant main effects of grammaticality (estimate = -3.049 (0.267), z = 11.38, p < .001) and distractor (estimate = 0.748 (0.102), z = 7.32, p < .001) were also found, revealing that grammatical conditions received more correct responses than ungrammatical conditions and responses to conditions without distractors were more accurate compared to conditions with distractors.

There was also a significant three-way interaction between group, grammaticality and distractor. Using nested contrasts, we tested grammaticality effects for each group in the distractor and no distractor conditions. This demonstrated significant grammaticality effects, with lower accuracy in ungrammatical conditions, for both groups in the two distractor/no distractor conditions. However, the effect was stronger for the L1 group. Specifically, for L1 speakers, the grammaticality effect was larger in the no distractor conditions (no distractor estimate = -5.26 (0.468), z = -11.24, p < .001; distractor estimate = -3.64 (0.371), z = -9.82, p < .001). It was also slightly larger for L2 speakers in the no distractor conditions (no distractor estimate = -1.88 (0.292), z = -6.44, p < .001; distractor estimate = -1.61 (0.269), z = -5.98, p < .001). Figure 9 depicts the pattern of the results.

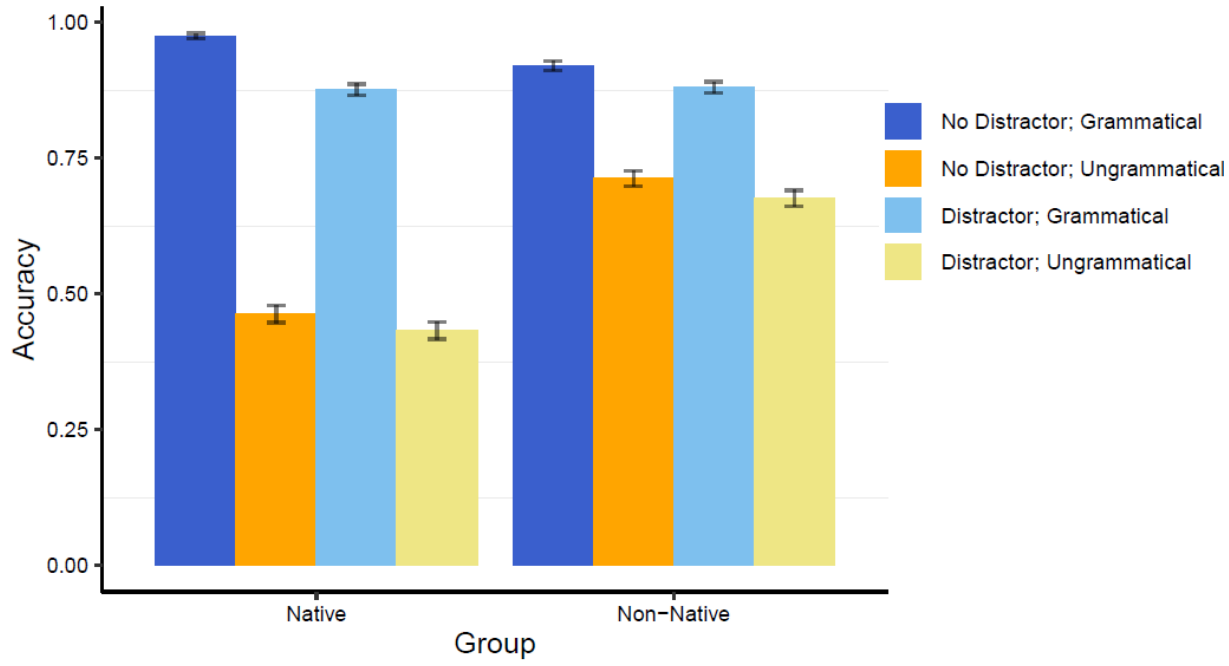


Figure 9. Accuracy Rates for Reflexive Dependencies in the Grammaticality Judgement Experiment in Study 2. Error Bars Represent Standard Errors.

4.3.1.2 Individual differences in L2 proficiency.

L2 proficiency had a significant main effect on L2 speakers' grammaticality judgements (estimate = 0.065 (0.010), z = 6.50, p < .001), such that accuracy rates increased as L2 proficiency increased. There was also a significant proficiency by distractor interaction (estimate = -0.024 (0.011), z = -2.07, p =.038). As shown in Figure 10, lower proficiency L2 speakers exhibited higher accuracy in the no distractor condition compared to the distractor conditions. However, this difference between conditions diminished as proficiency increased.

.

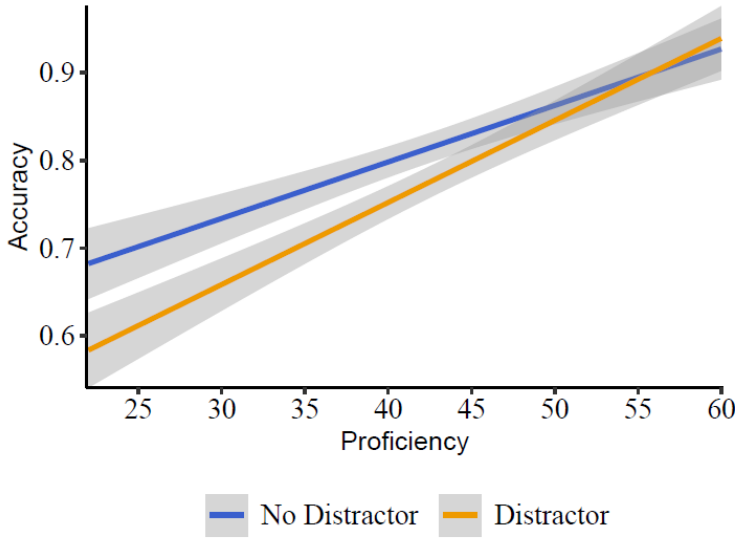


Figure 10. Effects of Proficiency and Distractor on L2 Speakers' Accuracy Rates in Study 2.

4.3.2 Self-paced reading.

4.3.2.1 Group analyses.

An average accuracy of 94% (range= 82-100%) was achieved by L1 speakers on comprehension questions, while L2 speakers achieved 93% accuracy (range=75-99%). A summary of mean reading times at the critical and spillover regions for both sets of experimental conditions is provided in Tables 13 and 14, respectively. The results obtained from the inferential statistics for both sets are also presented in Tables 15 and 16.

Table 13

Mean Reading Times (ms) by Condition for the First Set of Experimental Conditions in the Self-Paced Reading Experiment in Study 2.

	Critical Region		Spillover Region	
	L1 Speakers	L2 Speakers	L1 Speakers	L2 Speakers
Grammatical	424 (7)	776 (17)	393 (5)	526 (7)
Ungrammatical	450 (8)	789 (17)	463 (11)	580 (10)

Note. standard errors in parentheses.

Table 14

Mean Reading Times (ms) by Condition for the Second Set of Experimental Conditions in the Self-Paced Reading Experiment in Study 2.

	Critical Region		Spillover Region	
	L1 Speakers	L2 Speakers	L1 Speakers	L2 Speakers
Grammatical, Matched Distractor	453 (8)	757 (17)	401 (5)	566 (12)
Grammatical, Mismatched Distractor	468 (10)	722 (13)	406 (8)	581 (15)
Ungrammatical, Matched Distractor	479 (12)	807 (18)	463 (13)	622 (16)
Ungrammatical, Mismatched Distractor	465 (9)	792 (19)	443 (9)	623 (14)

Note. standard errors in parentheses.

Table 15

Statistical Analysis for the First Set of Experimental Conditions in the Self-Paced Reading Experiment in Study 2.

	Estimates (SE)	*t*	*p*
Group	0.41 (0.03)	15.08	<**0.001**
Grammaticality	0.05 (0.01)	5.77	<**0.001**
Region	-0.15 (0.02)	-7.20	<**0.001**
Group *Grammaticality	-0.03 (0.02)	-1.73	0.084

Group * Region	-0.28 (0.01)	-19.61	**<0.001**
Grammaticality *Region	0.06 (0.02)	3.04	**0.002**
Group*Grammaticality*Region	-0.01 (0.04)	-0.36	0.717

Note. significant effects ($p < 0.05$) are in bold.

Table 16

Statistical Analysis for the Second Set of Experimental Conditions in the Self-Paced Reading Experiment in Study 2.

	Estimates (SE)	*t*	*p*
Group	0.41 (0.03)	13.59	**<0.001**
Grammaticality	0.05 (0.01)	4.64	**<0.001**
Distractor	-0.01 (0.01)	-0.95	0.344
Region	-0.16 (0.01)	-22.19	**<0.001**
Group*Grammaticality	0.02 (0.01)	1.49	0.136
Group*Distractor	-0.01 (0.01)	-0.66	0.507
Grammaticality*Distractor	-0.01 (0.02)	-0.80	0.423
Group*Region	-0.19 (0.02)	-7.73	**<0.001**
Grammaticality*Region	0.04 (0.01)	3.86	**<0.001**
Distractor*Region	0.01 (0.01)	1.04	0.300
Group*Grammaticality*Distractor	0.01 (0.03)	0.28	0.780
Group*Grammaticality*Region	-0.04 (0.02)	-1.72	0.085
Group*Distractor*Region	0.04 (0.02)	1.96	0.051
Grammaticality*Distractor*Region	0.01 (0.02)	0.26	0.793
Group*Grammaticality*Distractor*Region	-0.00 (0.04)	-0.10	0.919

Note. significant effects ($p < 0.05$) are in bold.

As shown in Table 15, there were significant main effects of group and grammaticality in the baseline conditions, revealing that L2 speakers were slower than L1 speakers and reading times were longer in ungrammatical sentences relative to grammatical sentences (see Figure 11A). Additionally, a significant interaction between grammaticality and region was observed.

Nested contrasts demonstrated that grammaticality effects were significant at the spillover region but not at the critical region (critical region estimate = 0.025 (0.012), t = 2.05, p = .060; spillover estimate = 0.082 (0.017), t = 4.68, p < .001). However, the group by grammaticality interaction was not significant.

For the main set of experimental items, significant main effects of group and grammaticality were also observed. Grammaticality and region were also found to interact significantly. Nested contrasts revealed significant grammaticality effects at both regions, with larger effects observed at the spillover region (critical region estimate = 0.028 (0.008), t = 3.28, p = .001; spillover estimate = 0.068 (0.009), t = 7.26, p < .001), as illustrated in Figure 11B. Interestingly, no significant interaction was observed between group and grammaticality. Furthermore, no significant distractor effects were found.

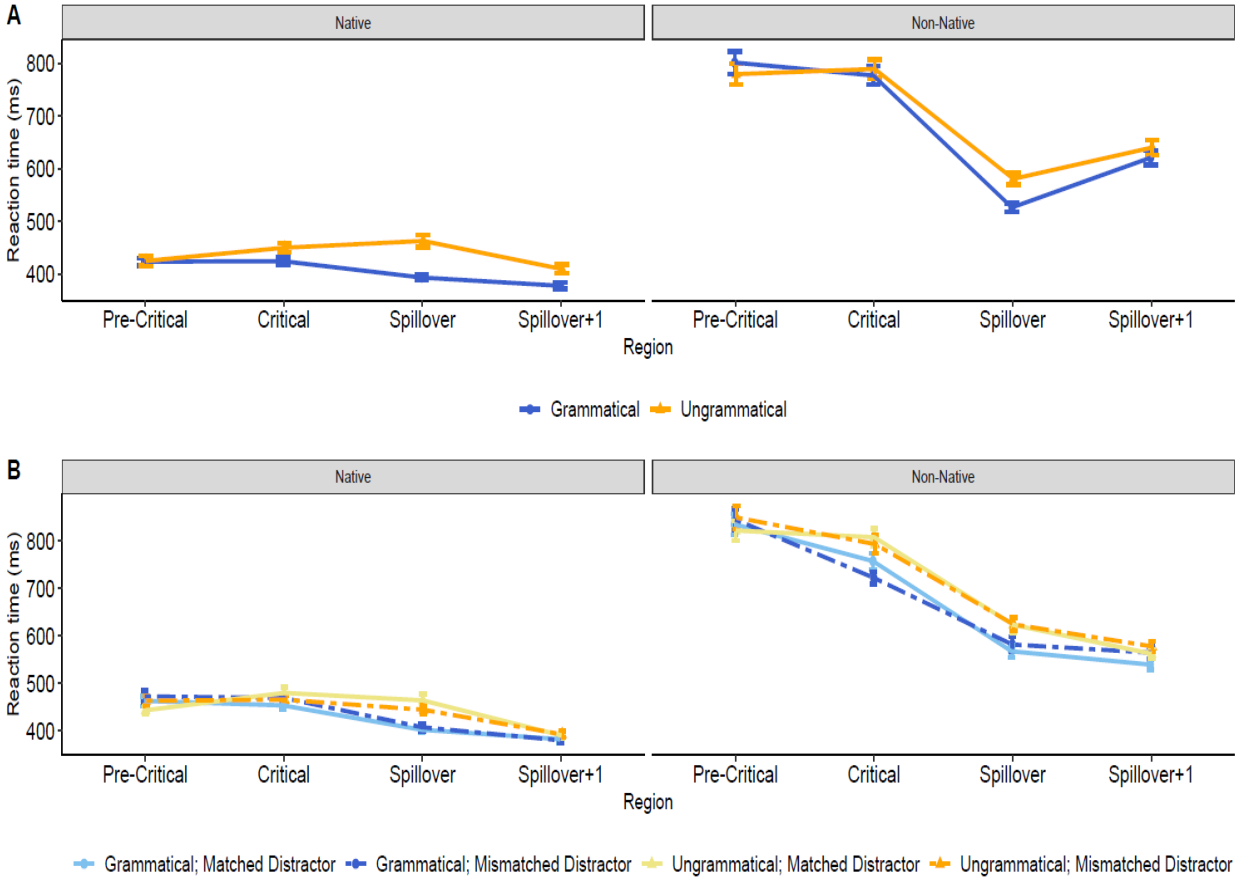


Figure 11. Reading Times for Reflexive Dependencies in the (A) baseline and (B) main experimental items in Study 2. Error Bars Represent Standard Errors.

4.3.2.2 Individual differences in L2 proficiency.

For the baseline conditions, the proficiency analysis revealed a significant interaction between grammaticality and proficiency (estimate = 0.004 (0.001), t = 3.15, p = .002), with lower proficiency L2 speakers showing larger grammaticality effects compared to higher proficiency L2 speakers, as shown in Figure 12.

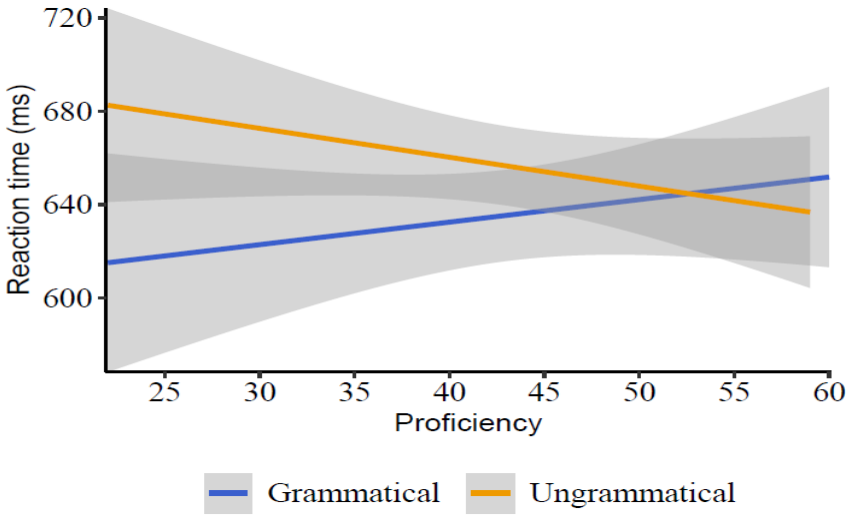


Figure 12. Interaction Effect between Proficiency and Grammaticality on L2 Speakers' Reading Times for the Baseline Conditions in Study 2.

For the main set of experimental items, proficiency analysis showed a significant three-way interaction between grammaticality, distractor and proficiency (estimate = -0.004 (0.002), t = -2.16, p = .041). As illustrated in Figure 13, reading times for grammatical conditions decreased as proficiency increased. Additionally, participants with higher proficiency levels appeared to be less affected by interference from matching distractors. In ungrammatical conditions, lower proficiency L2 speakers exhibited shorter reading times when the distractor matched the reflexive compared to when it did not, indicating facilitatory interference. Conversely, higher proficiency learners showed the opposite effect.

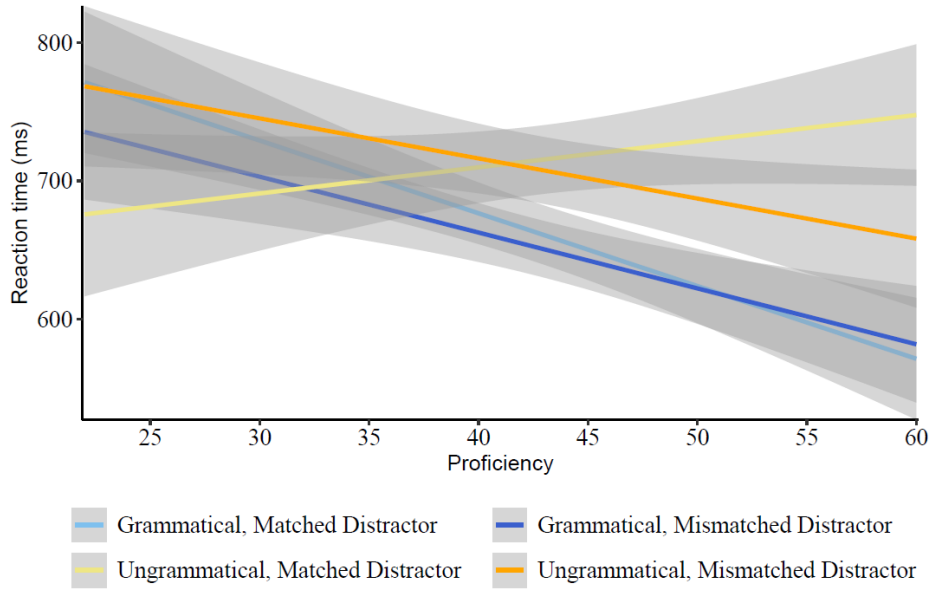


Figure 13. Interaction Effect Between Proficiency, Grammaticality and Distractor on L2 Speakers' Reading Times for the Main Set of Experimental Items in Study 2.

4.4 Discussion

4.4.1 Grammaticality judgements.

The offline grammaticality judgement data replicate those in Study 1 in finding accuracy rates higher in grammatical conditions than ungrammatical conditions for both groups, which might also indicate response bias. Despite this response bias, L2 speakers generally demonstrated knowledge of binding constraints. Indeed, we found L1 speakers were less accurate overall compared to L2 speakers. Numerically, this L1/L2 difference was largely carried by the ungrammatical conditions, which L1 speakers accepted roughly 50% of the time. This may suggest that L1 speakers occasionally interpreted the reflexive 'themselves' as a generic singular form, as has been sometimes observed with 'they' (Foertsch & Gernsbacher, 1997; Konnelly & Cowper, 2020; Speyer & Schleef, 2019), at least offline. As a result, L1 speakers may accept the conditions with singular antecedents. Though this possibility may have influenced both groups

when comparing their accuracy rates in ungrammatical conditions in Study 1 to Study 2, the influence is bigger on L1 speakers. This is possible as not all L2 speakers may possess knowledge of generic or gender-neutral pronouns, which have become increasingly popular and have been recently adopted by some L1 speakers (Ma, Wu, & Xu, 2022; Speyer & Schleef, 2019).

Proficiency analysis, however, additionally revealed that grammaticality and distractor effects on L2 speakers' judgements were moderated by proficiency. Specifically, accuracy increased with higher proficiency levels, and the slight difference in accuracy rates between distractor and no distractor conditions was most noticeable in L2 speakers with lower proficiency levels.

4.4.2 Self-paced reading.

L2 speakers exhibited similar grammaticality sensitivity to L1 speakers, whether in the baseline conditions or in the main set of experimental items. Both groups demonstrated clear grammaticality effects in both sets of experimental items. This is supported by the absence of a significant interactions between group and grammaticality. These findings suggest that the reduced grammaticality effects observed in L2 speakers compared to L1 speakers for S-V agreement compared to reflexives in Study 1 are related to the computation of agreement rather than the representation of the number feature itself. Note also that the grammaticality effects we observed in both the baseline and main experimental conditions suggest that, even if 'themselves' is sometimes treated as a generic reflexive, it is still also treated as plural. The manipulation of number in reflexives was not also found to induce any interference effects. These results, along with Study 1, will be discussed in more detail in the next chapter.

Although we did not find significant differences between L1 and L2 speakers in the main analysis of this study, our additional analysis of proficiency suggests that grammaticality sensitivity may be modulated by proficiency. However, unlike in Study 1 with reflexives, the size of grammaticality effects did not steeply get larger as proficiency levels rose. Since we did not

observe significant L1/L2 differences in the main analysis, we do not intend to draw strong conclusions here regarding individual differences in L2 speakers in Study 2.

CHAPTER FIVE

5. General Discussion and Conclusion

5.1 Overview

In this thesis, we sought to gather evidence from a large sample of Saudi-Arabic participants to investigate L2 speakers' sensitivity to syntactic constraints during S-V agreement and reflexive processing and determine whether they employ a qualitatively similar parsing mechanism to that of L1 speakers. Additionally, we examined whether similarity-based interference effects, which remain speculative in L2 processing literature, may influence L2 processing patterns, along with other individual differences factors. Since the susceptibility of these two linguistic dependencies to interference effects has been debated in L1 processing, testing them in L2 processing also provided an opportunity to contribute to this debate.

Our results generally indicate that L2 speakers' performance across different tasks aligned with that of L1 speakers, and any potential quantitative differences, that we did observe, might be driven by individual differences. Consistent interference effects were not observed, and L2 speakers were not found to be particularly more vulnerable to interference effects. We interpret the findings in more detail below.

5.2 S-V Agreement

Study 1 revealed L2 speakers' accurate untimed knowledge of S-V agreement constraints in line with previous offline findings (Chen et al., 2007; Lim & Christianson, 2014; Tanner et al., 2012). We also observed that L1 speakers were more affected by distractors than L2 speakers. The reason behind L1 speakers' judgement errors is not entirely clear. One possibility mentioned earlier is that L1 speakers' judgements were simply affected by the presence of two nouns. An alternative explanation discussed in Lee and Philips (2022) is that L2 speakers may show additional control over their L2 while L1 speakers may sacrifice accuracy for speed to get the task

done as fast as possible. We are cautious in drawing strong conclusions here however, given we did not manipulate the number properties of the distractor noun in our judgement task.

L2 speakers' reading times showed grammaticality effects at the spillover region, which is contrary to some previous studies that showed null effects (e.g., Chen et al., 2007; Jiang, 2004). The L2 grammaticality effects were however significantly smaller than those in L1 speakers. However, sensitivity to grammaticality was modulated by the level of L2 proficiency, supporting studies that found L1 and L2 processing of agreement is not fundamentally different (e.g., Cheng et al., 2022; Lim & Christianson, 2014; Tanner et al., 2012) and that proficiency plays a role in L2 attainment during agreement processing (Hopp, 2006; Jackson, 2008; Lim & Christianson, 2014; Sagarra & Herschensohn, 2010).

Contrary to expectations, the reading time patterns did not show any significant similarity-based interference effects usually found in L1 studies (Dillon et al., 2013; Jäger et al., 2020; Wagers et al., 2009) and also reported among L2 speakers in Tanner et al. (2012). Given that such effects have been reported previously, we are cautious about drawing strong conclusions regarding the lack of effects observed here. For present purposes, despite our large sample size, we found no significant evidence indicating that L2 speakers were more influenced by distractors compared to L1 speakers. Thus, in S-V agreement, our data suggest that L2 speakers are not more susceptible to interference than L1 speakers (cf. Cunnings, 2017), nor did we find evidence of shallow L2 processing (cf. Clahsen & Felser, 2006a, 2018). Instead, we argue that our agreement results suggest that L1 and L2 processing differ as a result of individual differences.

5.3 Reflexives

L2 speakers showed knowledge of Binding Principle A in our judgements tasks, consistent with the offline findings of the few existing studies (Felser et al., 2009; Felser & Cunnings, 2012). L1 speakers' judgements were also more affected by the presence of distractors than L2 speakers,

as was the case in S-V agreement, and we attribute it to the same potential issues described above in agreement.

Interestingly, our SPR results suggest that L2 speakers apply binding constraints during processing. Our L2 speakers demonstrated even larger grammaticality effects for gender mismatches than L1 speakers. This finding is contrary to previous studies (Felser et al., 2009; Felser & Cunnings, 2012) that suggested L2 reflexive processing is not initially constrained by Principle A, regardless of L2 speakers' L1. Across two experiments with gender and number manipulations, our results suggest L2 speakers rely more heavily on syntactic cues for reflexive resolution than gender, number or linear proximity. Our results may differ from previous studies for a number of reasons. One possibility might be that the distractor was more discourse prominent in the studies by Felser et al. and Felser and Cunnings, such that distractors may only influence L2 processing more than L1 processing when they are discourse prominent. In both of these studies, the distractor was initially introduced in a lead-in sentence before the critical sentence, which included the reflexive and the distractor in a discourse prominent subject position. However, in our experimental materials, the distractor was positioned in a non-discourse prominent position within a PP. Despite this, our results still nevertheless suggest that L2 speakers are not always more influenced by distractors than L1 speakers during anaphora resolution. The different findings between our experiments and previous studies could also be a result of the small sample-sizes in previous research compared to ours, and caution must be applied with small sample studies, which can overestimate effect sizes (Nicenboim & Vasishth, 2018; Vasishth & Gelman, 2021). This sheds light on the importance of replication and high-powered studies in L2 research.

The main effect of distractor observed with reflexive dependencies in Study 1 indicated shorter reading times in conditions with matching distractors. This effect was numerically larger in ungrammatical conditions, which we suggest is indicative of facilitatory interference in both

groups. No significant distractor effects were found in Study 2 however. Together, we interpret these results as suggesting that although syntactic cues may be more highly weighted during reflexive resolution than gender/number, reflexive resolution remains susceptible to interference (compare Cunnings & Sturt, 2014; Dillon et al., 2013; Jäger et al., 2020; Parker & Phillips, 2017).

5.4 Number in S-V Agreement and Reflexives

With regard to the reduced L2 sensitivity to grammaticality effects observed for S-V number agreement in Study 1, we did not observe significantly smaller grammatically effects for number violations during the processing of reflexives in Study 2. We interpret this as indicating that L2 speakers do not have difficulty with representing number per se, but rather the difficulty may lie in the processing of agreement. One possible explanation of this finding is that reflexive-antecedent dependencies generally have a stronger impact on sentence interpretation irrespective of the feature manipulated, gender or number, than agreement, and this may have additionally supported L2 processing. That is, reflexives are constituents that are underspecified in some way and hence require other elements in the same sentence to become fully interpretable (Felser, 2019). L2 speakers may pay more attention and focus more on feature (mis)match when it has a larger consequence on interpretation. However, we do not intend for this to imply that L2 speakers focus on semantics over syntax, as our L2 speakers generally demonstrated understanding of structural constraints on dependency formation during processing. There were no significant differences between them and L1 speakers in Study 2, and they exhibited even larger grammaticality effects compared to L1 speakers in Study 1. Additionally, this does not imply the ease of processing reflexives, as the processing of reflexives specifically requires considering the hierarchical relationships between constituents and promptly applying the constraint that prevents illicit licensers from being considered (Felser, 2019). Moreover, note also that L2 grammaticality effects interacted with proficiency in S-V agreement, suggesting a more native-like performance

in higher proficiency L2 speakers. This indicates that the L1/L2 difference in S-V agreement is influenced by individual differences. Below, we discuss the broader implications of the findings.

5.5 Resolving Linguistic Dependencies in L1 and L2 Processing

Our findings provided insight into the similarities and differences between native and non-native processing of syntactically mediated dependencies. As discussed in Chapter 2, the Shallow Structure Hypothesis (Clahsen & Felser, 2006a, 2018) argued that L2 speakers, even highly proficient ones, follow a shallow processing route during real-time sentence comprehension that is guided more by lexical semantics, pragmatic and discourse information than the grammar due to their reduced syntactic parsing ability, in contrast to L1 speakers who perform full syntactic analysis. Our offline data showed that L2 speakers possessed adequate grammatical knowledge of S-V agreement and reflexive-antecedent dependencies. The L2 grammar therefore had the required syntactic information to process the linguistic phenomena under consideration, but online data would reveal whether the relevant syntactic constraints could be successfully applied in real-time. In our online experiments, we manipulated sentence grammaticality and the properties of the distractor to examine whether L2 speakers underuse syntactic constraints and tolerate agreement and Binding Condition A violations or form dependencies with matching illicit elements if their computation is more derived by meaning as argued by the SSH, compared to L1 speakers. The results, however, demonstrated that grammaticality effects and the utilisation of syntactic constraints were evident in L2 speakers' processing patterns in both Study 1 and 2, which is not compatible with the Shallow Structure Hypothesis (Clahsen & Felser, 2006a, 2018). L2 speakers' reading times would indicate shallow processing if there were no demonstrated grammaticality effects or if reading times were longer for grammatical sentences with mismatching distractors and shorter for ungrammatical sentences with matching distractors. This was not the case, however.

More recently, the difficulty in processing discontinuous dependencies has also been predicted to vary based on the nature of the search process involved. Specifically, it was expected that L2 speakers' reduced sensitivity to syntactic constraints would affect backward-looking dependencies more considerably than forward-looking dependencies (Felser, 2015). Felser argued that L2 speakers, even with intermediate proficiency levels, might show sensitivity to syntactic constraints, including those that are not instantiated in their L1, when processing forward-looking dependencies. However, it was expected that they would violate constraints for backward-looking dependencies, even if those constraints were instantiated in their L1 in the same way (Felser, 2015). Accordingly, the processing of reflexive-antecedent dependencies, in particular, was anticipated to pose a greater challenge for L2 speakers, since it is more clearly a backward-looking dependency than S-V agreement, which might be considered a hybrid type of dependency that potentially involves both forward-looking (predictive) and backward-looking searches (Felser, 2015). Our findings however, revealed that L2 speakers demonstrated native-like processing performance for reflexive-antecedent dependencies, as a backward-looking dependency, which stands against what has been argued for.

Other competing accounts of L2 sentence processing, on the other hand, argued that L2 speakers construct syntactic parses similarly to L1 speakers, and potential differences can be due to other factors (Cunnings, 2017; Hopp, 2014, 2015, 2018; McDonald, 2006). Cunnings (2017) proposed that L2 speakers might be more vulnerable than L1 speakers to similarity-based interference effects that may result from the memory retrieval process involved during dependency formation. Our experimental manipulations allowed us to test this account. We found L2 speakers make use of syntactic cues like L1 speakers during real-time processing, but we did not find them more susceptible to interference effects than L1 speakers. Interference effects, specifically facilitatory interference, as supported by the faster reading times for ungrammatical sentences with gender-matching distractors compared to those with mismatching distractors,

could only be observed in reflexives with gender manipulation, similarly in both groups. As such, our findings do not conform to Cunnings' (2017) proposal, but they show that L2 speakers' processing can undergo the same effects associated with the memory retrieval mechanism as L1 speakers, even though these effects were not consistently found.

Finding evidence of facilitatory interference during the processing of reflexives contributes to the debate in the L1 processing literature regarding the nature of linguistic cues utilised as retrieval cues when processing reflexive-antecedent compared to S-V agreement dependencies (Dillon et al., 2013; Jäger et al., 2020; Parker & Phillips, 2017; Phillips, Wagers, & Lau, 2011). While cue-based parsing models suggest a single memory retrieval mechanism employed for resolving different linguistic dependencies, implying similar potential for interference effects (Lewis & Vasishth, 2005; Lewis et al., 2006; McElree, 2000; McElree et al., 2003), some argued that S-V agreement and reflexives-antecedent dependencies do not recruit surface agreement features as retrieval cues in the same way. Consequently, interference effects are typically observed for S-V agreement but not for reflexives (Dillon et al., 2013; Parker & Phillips, 2017). The latter assumes that surface agreement features are not involved as retrieval cues in forming reflexive binding dependencies since they are not directly implicated in the grammatical binding relation, whereas S-V agreement dependencies rely on feature content agreement between two elements, and so they naturally serve as retrieval cues for constructing agreement dependencies (Dillon et al., 2013). Though we unexpectedly did not find evidence of interference in S-V agreement, we did observe it in reflexives in both groups. Our findings, therefore, suggest that processing reflexive-antecedent dependencies can be susceptible to interference effects, which indicates that surface agreement features could also be used during reflexive processing, and sensitivity to different linguistic features may not necessarily differ across distinct grammatical dependencies. However, we are cautious in drawing strong

conclusions based on our results regarding the absence of interference in S-V agreement, given its usual occurrence in prior studies.

Furthermore, the fact that we did not find symmetrical 'attraction' effects in grammatical and ungrammatical conditions that might arise as suggested from a faulty or unclear representation of the subject goes against representational accounts (Bock & Eberhard, 1993; Eberhard et al., 2005; Eberhard, 1997). Nevertheless, one might argue that the effect observed for reflexives in ungrammatical sentences could potentially be explained by both representational accounts and cue-based memory retrieval models. However, the effect observed is more in line with the cue-based retrieval model, as representational accounts anticipate similar effects in both grammatical and ungrammatical conditions (Eberhard, 1997). Not finding the same size of effects across grammatical and ungrammatical conditions has essentially been taken in the literature as supporting evidence for the cue-based memory retrieval model (Wagers et al., 2009). In addition, interference effects that might result from the cue-based memory retrieval mechanism have rarely been reported in the literature for both grammatical and ungrammatical conditions within the same study. Further, according to the cue-based memory retrieval model, the effect predicted in grammatical conditions is relatively smaller and less likely to occur compared to the effect predicted in ungrammatical sentences (Jäger et al., 2017, 2020; Nicenboim & Vasishth, 2018). Therefore, we interpreted the asymmetrical attraction effect observed in both groups as indicative of facilitatory interference.

Apart from the parsing mechanism involved during dependency formation, L1/L2 processing differences have alternatively been thought to be largely defined by individual differences (Hopp, 2014, 2015, 2018; McDonald, 2006), as discussed earlier. Proficiency, for instance, has been considered a crucial factor that affects L2 processing in that lower proficiency levels may mask target-like processing among L2 speakers (Hopp, 2006). By testing proficiency differences in a large sample of L2 speakers, we found proficiency affects L2 speakers' accuracy

and processing at a significant level. This suggests that linguistic representations develop, and target-like processing occurs as proficiency advances. More specifically, our findings showed that L1/L2 syntactic processing resemble each other and the L1/L2 processing differences, as particularly observed in S-V agreement, are quantitative in nature and can be ameliorated at higher levels of proficiency. This also accords with previous agreement studies that found sensitivity to morphosyntactic agreement violations specifically is largely determined by the level of proficiency (e.g., Coughlin & Tremblay, 2012).

L2 processing has also been suggested to be influenced by the automaticity of lexical processing (Hopp, 2018). As this entails efficient lexical access and retrieval that might precede structure-building processes, difficulties or delays in lexical processing have been assumed to obscure the effects of syntactic structure and target-like processing (Hopp, 2018). Our findings, however, did not show reliable effects of lexical automaticity on L2 speakers' processing, especially during processing S-V agreement dependencies, where reduced grammaticality effects have been observed. It could be, however, that lexical processing does not influence different linguistic dependencies in the same way, or that individual differences in lexical automaticity cannot explain all difficulties with syntactic processing.

Though our main interest lies in examining the working memory operations that contribute to sentence comprehension, namely the effects of the cue-based memory retrieval mechanism, rather than the capacity-based effects as indicated earlier, the effects of the latter can still be considered through the linguistic materials, but not based on differences in working memory capacity across participants. As there was no clear preference for the linearly closest noun to the dependent element, there is no evidence for capacity-based effects, though a distance manipulation might technically be needed to confirm this conjecture. Among different factors, differences in proficiency seem to have played a vital role in L2 processing of S-V agreement and reflexive-antecedent dependencies.

In sum, across studies, we did not find evidence that L2 speakers are more susceptible to interference than L1 speakers (Cunnings, 2017). We also did not find evidence of shallow L2 parsing (cf. Clahsen & Fesler, 2006a, 2018). Instead, our results suggest L2 processing can be target-like, resembling L1 processing, with individual differences explaining potential differences.

5.6 Limitations and Future Directions

The scope of this thesis revolves around the similarities and differences in L1/L2 processing, particularly focusing on memory operations involved in sentence processing. Our specific aim was to investigate the real-time effects of the cue-based memory retrieval mechanism on L2 speakers' reading times. The off-line grammaticality judgement tasks were intended to assess L2 speakers' grammatical knowledge of the linguistic phenomena under investigation. In the grammaticality judgement tasks, we tested grammaticality in dependencies that contained only one noun phrase to examine the grammatical representations of S-V agreement and reflexives by L2 speakers, in basic terms. We also tested grammaticality in discontinuous dependencies that contained two noun phrases to confirm that they accurately form a dependency with the target dependent element and not with any other element. Participants demonstrated a satisfactory understanding of the relevant grammatical constraints. However, it was somewhat unexpected to find participants' accuracy rates, particularly L1 speakers', largely higher in conditions that contained only one noun phrase compared to conditions with two noun phrases. As we did not manipulate the distractor's properties as in the online task, it was difficult to assess precisely the factor that may have created this difference in accuracy rates. That is, including conditions with the same length or number of noun phrases but manipulating the similarity of the distractor could have aided in providing a more nuanced explanation. For instance, if participants' accuracy rates are significantly lower in longer conditions, regardless of the distractor type, compared to conditions containing only one noun phrase, this might suggest capacity-based

limitations effects. However, variations in accuracy rates depending on the type of the distractor may indicate the presence of similarity-based interference effects. The lack of such data that might be required to tease apart these possibilities has made it difficult to be conclusive.

Alternatively, the role of memory capacity could have been tackled better by including a measure of individual differences in readers' working memory capacity. Accordingly, the absence of a significant interaction between working memory capacity and accuracy rates, for example, can rule out the possibility of capacity-related effects. The observed difference in accuracy rates then may just indicate that sustained attention is better for shorter sentences. Compared to L1 speakers, L2 speakers may, however, employ an additional monitoring mechanism, as suggested by Lee and Philips (2022), allowing them to filter explicit judgements and exhibit greater control over their L2. We also mentioned that capacity-based effects on online processing can be considered through the linguistic materials. Our findings did not show evidence of such effects by forming dependencies, for example, with the linearly closest noun phrase, which may also indicate a type of shallow processing. However, if we had observed such effects, particularly among L2 speakers, it would have been difficult to determine whether they resulted from a capacity-based limitation or shallow processing. Therefore, including a measure of individual differences in readers' working memory capacity would also be helpful for gaining a better understanding of online data and controlling potential confounding effects.

It is also worth mentioning that the pattern of findings offers insights into the importance of investigating different linguistic phenomena within the same large-scale controlled study. To increase certainty in linguistic work, further large-scale controlled studies are needed to systematically examine other linguistic phenomena. Replicating the study in other L2 languages would also help validate the findings. Additionally, the observation that L2 speakers exhibited larger grammaticality effects for reflexives with gender manipulation relative to S-V agreement may suggest that testing gender along with number in S-V agreement could be a possible

progression of this work. If L2 speakers have specific difficulties with agreement, they should also be less sensitive to gender during S-V agreement processing than L1 speakers. Since gender agreement is not marked in English, this can be an avenue for future research in other L2 languages. Although a few empirical investigations have been conducted on gender agreement within the nominal domain (between determiner-noun-adjective combinations) in L2 Spanish using event-related potentials (Gabriele et al., 2021, 2013), further work using other experimental measures and L2 languages examining also verbal gender agreement (between subject and verb) may establish a greater degree of accuracy on this matter.

5.7 Practical Implications

Examining L2 processing of linguistic dependencies in light of L2 processing theoretical accounts provides insights into the potential factors and cognitive processes that may impact language comprehension. This, in turn, contributes to deepening our understanding of language teaching and learning, which might not be as straightforward as might be assumed. Moreover, investigating how L2 speakers apply linguistic knowledge in real time, rather than solely studying their knowledge of specific aspects of the target language, not only demonstrates their language competence but also shows their ability to use it in a contextual setting.

If L2 speakers have difficulties in real time comprehension, this has important implications for teaching. From a psycholinguistic perspective on language learning, teachers should perceive learning as an individual process wherein knowledge and skills are acquired and improved through exposure to L2 comprehensible input and negotiation (Carlos, 2008). By manipulating the input learners encounter through interactive tasks, teachers can increase the chances of comprehensible input and enhance language development (Long, 1983). Teachers can employ various interactional modifications, such as clarification requests, comprehension questions and confirmation checks to ensure accurate input comprehension and facilitate

negotiation for meaning. This, in turn, provides opportunities to notice form, meaning and language use (Carlos, 2008). Consequently, teachers can assess whether learners' language processing is influenced by interference effects, grammatical or lexical knowledge. In other words, the current study helps to predict the challenges learners encounter during language use and offers new directions for error analysis, which can effectively facilitate the creation of a productive learning environment.

In line with this, the study also sheds light on the importance of using psycholinguistic methods in L2 classrooms to assess language development. Relying on paper-and-pencil tests, which evaluate explicit knowledge of grammar or vocabulary based on taught material, may not accurately measure the extent to which students have fully assimilated that knowledge (Marinis and Cunnings, 2018). This discrepancy might explain the variations observed between learners' performance in such written tests and their ability to use that knowledge for comprehension or communication in real life situations. Therefore, utilising psycholinguistic methods to assess language comprehension and production holds practical value. Consistently evaluating students' language development through performance-based examinations integrated with standardized written tests can help teachers modify teaching strategies based on students' progress. Teachers can promptly intervene by incorporating meaningful learning activities, adjusting ineffective strategies, allocating less time to completely mastered things or providing new challenges. For example, the present study revealed that L2 speakers have some issues or difficulties when processing subject-verb agreement as compared to reflexives, this may partly suggest that they may have not really internalised relevant rules or that they have not been given enough room to practice what they have been taught. Consequently, additional time should be allocated for teaching subject-verb agreement.

Examining how individual differences can modulate L2 speakers' performance has also significant implications for teaching. Given that learners vary in proficiency, lexical knowledge,

processing speed, motivation, and other aspects, it is crucial for teachers to acknowledge these differences and address the issue of classroom heterogeneity. A better understanding of individual differences helps teachers to design learning experiences that cater to the specific needs of each student. For instance, if students exhibit variations in proficiency levels or vocabulary knowledge, grouping them in complementary ways encourage peer learning, or additional support can be provided in other settings when necessary. Moreover, accommodating the diverse characteristics of learners may involve offering individualised instruction by adjusting the level, pace, learning tasks or teaching styles to engage students and meet their specific needs. Therefore, to enhance the learning experience, practitioners need to develop a better understanding of the factors contributing to individual variability in language learning and adapt their teaching approaches accordingly.

In sum, this section has sought to discuss some of the study's implications for teaching and learning, addressing generally some factors that may assist in developing pedagogical practices that foster language learning.

5.8 Conclusion

Considering the discrepancy in L2 processing research regarding L1/L2 processing similarities and differences during real-time dependency formation, we conducted large-scale experiments testing two linguistic dependencies with the same group of L2 speakers. Through examining S-V agreement and reflexives, our results revealed that applying syntactic constraints during processing is not always more problematic in L2 processing than in L1 processing. Though restricting L1/L2 processing differences to a single source can be difficult, as different linguistic dependencies may involve distinct constraints, our findings generally did not indicate that L1/L2 processing is fundamentally different in nature, but rather moderated by individual differences in certain contexts. That is, proficiency variations may influence target-like grammatical parsing in

L2 processing but do not constrain it, particularly for higher proficiency L2 speakers, supporting the notion of continuity between native and non-native processing. Therefore, we argue that investigating different linguistic dependencies within the same L2 group, while controlling for various variables, cannot only enhance our understanding of L2 processing, but also the accuracy and generalisability of findings.

References

AlBzour, N. N. (2015). Arabic Anaphora: Discourse Subduing Morpho-Syntactic Interaction. *English Language and Literature Studies*, *5*(3), 38–50. https://doi.org/10.5539/ells.v5n3p38

Aldwayan, S., Fiorentino, R., & Gabriele, A. (2010). Evidence of syntactic constraints in the processing of wh-movement: A study of Najdi Arabic learners of English. In B. VanPatten & J. Jegerski (Eds.), *Research in Second Language Processing and Parsing* (pp. 65–86). Amsterdam: John Benjamins. https://doi.org/10.1075/LALD.53.03ALD

Alemán Bañón, J., Fiorentino, R., & Gabriele, A. (2018). Using event-related potentials to track morphosyntactic development in second language learners: The processing of number and gender agreement in Spanish. *PloS One*, *13*(7), 1–35. https://doi.org/10.1371/journal.pone.0200791

Alemán Bañón, J., & Rothman, J. (2016). The role of morphological markedness in the processing of number and gender agreement in Spanish: an event-related potential investigation. *Language, Cognition and Neuroscience*, *31*(10), 1273–1298. https://doi.org/10.1080/23273798.2016.1218032

AlKafri, A. (2013). *Interpretation of English Reflexives by Child and Adult L2 Learners.* (PhD Thesis, Newcastle University.). Retrieved from http://theses.ncl.ac.uk/jspui/handle/10443/1845

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An Integrated Theory of the Mind. *Psychological Review*, *111*(4), 1036–1060. https://doi.org/10.1037/0033-295X.111.4.1036

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/J.JML.2007.12.005

Baddeley, A. (2007). *Working Memory, Thought, and Action*. Oxford, UK: Oxford University Press.

Badecker, W., & Straub, K. (2002). The Processing Role of Structural Constraints on the Interpretation of Pronouns and Anaphors. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(4), 748–769. https://doi.org/10.1037/0278-7393.28.4.748

Baker, M. (2008). *The Syntax of Agreement and Concord*. Cambridge, UK: Cambridge University Press.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, *17*(1), 137–153. https://doi.org/10.1162/0898929052880101

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/J.JML.2012.11.001

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Benduhaish, S. O. A. (2018). *Features Between Form and Interpretation : The Effect of Animacy , Gender , and Number on Verbal Agreement* (PhD Thesis, University of Leeds). Retrieved from https://etheses.whiterose.ac.uk/22245/

Benmamoun, E. (1992). *Functional and inflectional morphology problems of projection, representation and derivation* (PhD Thesis, University of Southern California).

Bergen, L., & Gibson, E. (2012). Agreement errors as rational encoding errors. *Poster Presented at the 25th CUNY Conference on Human Sentence Processing.* New York, US.

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in english number agreement. *Language and Cognitive Processes*, *8*(1), 57–99. https://doi.org/10.1080/01690969308406949

Bock, K., & Miller, C. A. (1991). Broken Agreement. *Cognitive Psychology*, *23*(1), 45–93. https://doi.org/10.1016/0010-0285(91)90003-7

Carlos, M. S. C. (2008). Psycho-linguistic and socio-cultural approaches to language learning: A never ending debate. *Colombian Applied Linguistics Journal*, *10*(1), 142–154. http://www.scielo.org.co/pdf/calj/n10/n10a08.pdf

Carstens, V. (2000). Concord in Minimalist Theory. *Linguistic Inquiry*, *31*(2), 319–355. https://doi.org/10.1162/002438900554370

Chemla, E., Cummins, C., & Singh, R. (2017). Training and timing local scalar enrichments under global pragmatic pressures. *Journal of Semantics*, *34*(1), 107–126. https://doi.org/10.1093/jos/ffw006

Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject – verb agreement in L2 learning∗. *Bilingualism: Language and Cognition*, *10*(2), 161–174. https://doi.org/10.1017/S136672890700291X

Cheng, Y., Cunnings, I., Miller, D., & Rothman, J. (2021). Double-number marking matters for both L1 and L2 processing of nonlocal agreement similarly: an ERP investigation. *Studies*

in Second Language Acquisition, *44*(5), 1309–1329. https://doi.org/10.1017/S0272263121000772

Cheng, Y., Rothman, J., & Cunnings, I. (2021). Parsing preferences and individual differences in nonnative sentence processing: Evidence from eye movements. *Applied Psycholinguistics*, *42*(1), 129–151. https://doi.org/10.1017/S014271642000065X

Cheng, Y., Rothman, J., & Cunnings, I. (2022). Determiner-Number Specification and Non-Local Agreement Computation in L1 and L2 Processing. *Journal of Psycholinguistic Research*, *51*, 847–863. https://doi.org/10.1007/s10936-022-09864-w

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht, The Netherlands: Fori.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use.* Westport, Connecticut: USA: Greenwood Publishing Group.

Chomsky, N. (1995). *The Minimalist program*. Cambridge: MA: The MIT Press.

Clackson, K., & Heyer, V. (2014). Reflexive anaphor resolution in spoken language comprehension: Structural constraints and beyond. *Frontiers in Psychology*, *5*(Article 904), 1–9. https://doi.org/10.3389/fpsyg.2014.00904

Clahsen, H., & Felse, C. (2006). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, *27*(1), 107–126. https://doi.org/10.1017/S0142716406060206

Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*(1), 3–42. https://doi.org/10.1017/S0142716406060024

Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Science*, *10*(12), 564–570. https://doi.org/10.1016/j.tics.2006.10.002

Clahsen, H., & Felser, C. (2018). Some notes on the Shallow Structure Hypothesis. *Studies in*

Second Language Acquisition, *40*(3), 693–706. https://doi.org/10.1017/S0272263117000250

Corbett, G. G. (2006). *Agreement*. Cambridge, UK: Cambridge University Press.

Cotter, B. T., & Ferreira, F. (2022). The Relationship Between Working Memory Capacity, Bilingualism, and Ambiguous Relative Clause Attachment. *PsyArXiv*. https://doi.org/10.31234/OSF.IO/Y3GM4

Coughlin, C. E., & Tremblay, A. (2012). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics*, *34*(3), 615–646. https://doi.org/10.1017/S0142716411000890

Cowan, N. (1988). Evolving Conceptions of Memory Storage, Selective Attention, and Their Mutual Constraints Within the Human Information-Processing System. *Psychological Bulletin*, *104*(2), 163–191. https://doi.org/10.1037/0033-2909.104.2.163

Cunnings, I. (2017). Parsing and Working memory in bilingual sentence Processing. *Bilingualism: Language and Cognition*, *20*(4), 659–678. https://doi.org/10.1017/S1366728916000675

Cunnings, I. (2022). Interference-based and capacity-based approaches to working memory in second language sentence processing. In T. Leal, E. Shimanskaya, & C. A. Isabelli (Eds.), *Generative SLA in the age of Minimalism: Features, interfaces, and beyond. Selected proceedings of the 15th Generative Approaches to Second Language Acquisition Conference*. (pp. 243–264). Amsterdam: John Benjamins https://doi.org/10.1075/LALD.67.10CUN

Cunnings, I., & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language*, *75*, 117–139. https://doi.org/10.1016/j.jml.2014.05.006

Cunnings, I., & Sturt, P. (2018). Coargumenthood and the processing of pronouns. *Language, Cognition and Neuroscience*, *33*(10), 1235–1251. https://doi.org/10.1080/23273798.2018.1465188

Dallas, A., & Kaan, E. (2008). Second Language Processing of Filler-Gap Dependencies by Late Learners. *Linguistics and Language Compass*, *2*(3), 372–388. https://doi.org/10.1111/j.1749-818X.2008.00056.x

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/s0022-5371(80)90312-6

Dillon, B., Clifton, C., & Frazier, L. (2013). Pushed aside: Parentheticals, memory and processing. *Language, Cognition and Neuroscience*, *29*(4), 483–498. https://doi.org/10.1080/01690965.2013.866684

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103. https://doi.org/10.1016/j.jml.2013.04.003

Dowens, M. G., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of Cognitive Neuroscience*, *22*(8), 1870–1887. https://doi.org/10.1162/JOCN.2009.21304

Drummond, A. (2013). *Ibex Farm.* Retrieved from http://spellout.net/ibexfarm/.

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *10*(1), 101–116. https://doi.org/10.1017/S1366728906002847

Eberhard, K. M. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language*, *36*(2), 147–164. https://doi.org/10.1006/jmla.1996.2484

Eberhard, K. M., Cutting, J. C., & Bock, K. (2005). Making Syntax of Sense: Number Agreement in Sentence Production. *Psychological Review*, *112*(3), 531–559. https://doi.org/10.1037/0033-295X.112.3.531

Engelmann, F. (2016). *Toward an Integrated Model of Sentence Processing in Reading* (PhD Thesis, University of Potsdam). Retrieved from https://www.ling.uni-potsdam.de/~vasishth/courses/pdfs/SP-EM-Model.pdf

Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Farmer, T. A., Misyak, J. B., & Christiansen, M. H. (2012). Individual Differences in Sentence Processing. In M. J. Spivey, K. McRae, & M. F. Joanisse (Eds.), *The Cambridge Handbook of Psycholinguistics.* (pp. 353–364). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139029377.024

Fehri, A. F. (1993). *Issues in the Structure of Arabic Clauses and Words (Vol. 29)*. Dordrecht: Springer Science and Business media. https://doi.org/10.1007/978-94-017-1986-5

Felser, C. (2015). Native vs. Non-native Processing of Discontinuous Dependencies. *Second Language*, *14*, 5–19. https://doi.org/10.11431/secondlanguage.14.0_5

Felser, C. (2019). Structure-sensitive constraints in non-native sentence processing. *Journal of the European Second Language Association*, *3*(1), 12–22. https://doi.org/10.22599/jesla.52

Felser, C., & Cunnings, I. (2012). Processing reflexives in a second language : The timing of structural and discourse-level constraints. *Applied Psycholinguistics*, *33*(3), 571–603.

https://doi.org/10.1017/S0142716411000488

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, *23*(1), 9–36. https://doi.org/10.1177/0267658307071600

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*(3), 453–489. https://doi.org/10.1017/S0142716403000237

Felser, C., Sato, M., & Bertenshaw, N. (2009). The on-line application of binding principle A in english as a second language. *Bilingualism: Language and Cognition*, *12*(4), 485–502. https://doi.org/10.1017/S1366728909990228

Foertsch, J., & Gernsbacher, M. A. (1997). In search of gender nertrality: Is singular they a cognitively efficient substitute for generic he? *Psychological Science*, *8*(2), 106–111. https://doi.org/10.1111/J.1467-9280.1997.TB00691.X

Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics*, *32*(1), 187–220. https://doi.org/10.1017/S0142716410000342

Foraker, S., & McElree, B. (2011). Comprehension of linguistic dependencies: Speed-accuracy tradeoff evidence for direct-access retrieval from memory. *Linguistics and Language Compass*, *5*(11), 764–783. https://doi.org/10.1111/j.1749-818X.2011.00313.x

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity*. *Bilingualism: Language and Cognition*, *14*(3), 379–399. https://doi.org/10.1017/S136672891000012X

Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, *17*(3), 213–247. https://doi.org/10.1177/026765830101700301

Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender*. Amsterdam: John Benjamins.

Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, *17*(4), 371–404. https://doi.org/10.1080/01690960143000254

Gabriele, A., Bañón, J. A., Hoffman, L., Covey, L., Rossomondo, A., & Fiorentino, R. (2021). Examining Variability in the Processing of Agreement in Novice Learners: Evidence From Event-Related Potentials. *Journal of Experimental Psychology: Learning Memory and Cognition*, *47*(7), 1106–1140. https://doi.org/10.1037/XLM0000983

Gabriele, A., Fiorentino, R., & Alemán Bañón, J. (2013). Examining second language development using event-related potentials. *Linguistic Approaches to Bilingualism*, *3*(2), 213–232. https://doi.org/10.1075/lab.3.2.04gab

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency Drives Lexical Access in Reading but Not in Speaking: The Frequency-Lag Hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209. https://doi.org/10.1037/a0022256

Haegeman, L. M. V. (1994). *Introduction to government and binding theory*. Oxford: Blackwell.

Hammerly, C., Staub, A., & Dillon, B. (2019). The grammaticality asymmetry in agreement attraction reflects response bias: Experimental and modeling evidence. *Cognitive*

Psychology, *110*, 70–104. https://doi.org/10.1016/j.cogpsych.2019.01.001

Hawkins, R. (2001). *Second language syntax: A generative introduction*. Oxford: Blackwell. http://repository.essex.ac.uk/id/eprint/11581

Hawkins, R., & Chan, C. Y. H. (1997). The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis.' *Second Language Research*, *13*(3), 187–226. https://doi.org/10.1191/026765897671476153

Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In P. Prévost & J. Paradis (Eds.), *The acquisition of French in different contexts* (pp. 175–205). Amsterdam: John Benjamins.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, *22*(3), 369–397. https://doi.org/10.1191/0267658306sr272oa

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*(4), 901–931. https://doi.org/10.1016/j.lingua.2009.06.004

Hopp, H. (2014). Working Memory Effects in the L2 Processing of Ambiguous Relative Clauses. *Language Acquisition*, *21*(3), 250–278. https://doi.org/10.1080/10489223.2014.892943

Hopp, H. (2015). Individual differences in the second language processing of object–subject ambiguities. *Applied Psycholinguistics*, *36*(2), 129–173. https://doi.org/10.1017/S0142716413000180

Hopp, H. (2018). The bilingual mental lexicon in L2 sentence processing. *Second Language*,

17, 5–27. https://www.jstage.jst.go.jp/article/secondlanguage/17/0/17_5/_pdf/-char/ja

Jackson, C. (2008). Proficiency level and the interaction of lexical and morphosyntactic information during L2 sentence processing. *Language Learning*, *58*(4), 875–909. https://doi.org/10.1111/j.1467-9922.2008.00481.x

Jäger, L. A., Engelmann, F., & Vasishth, S. (2015). Retrieval interference in reflexive processing: experimental evidence from Mandarin, and computational modeling. *Frontiers in Psychology*, *6*(Article 617), 1–24. https://doi.org/10.3389/fpsyg.2015.00617

Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, *94*, 316–339. https://doi.org/10.1016/j.jml.2017.01.004

Jäger, L. A., Mertzen, D., Van Dyke, J. A., & Vasishth, S. (2020). Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language*, *111*(104063), 1–21. https://doi.org/10.1016/j.jml.2019.104063

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*(4), 603–634. https://doi.org/10.1017/s0142716404001298

Juffs, A. (2004). Representation, Processing and Working Memory in a Second Language. *Transactions of the Philological Society*, *102*(2), 199–225. https://doi.org/10.1111/j.0079-1636.2004.00135.x

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*(1), 122–149. https://doi.org/10.1037/0033-295X.99.1.122

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative

Spanish: An eye-movement investigation. *Language Learning*, *59*(3), 503–535. https://doi.org/10.1111/j.1467-9922.2009.00516.x

Keating, G. D. (2010). The effects of linear distance and working memory on the processing of gender agreement in Spanish. In B. Vanpatten & J. Jegerski (Eds.), *Research in Second Language Processing and Parsing* (pp. 113–134). Amsterdam: John Benjamins. https://doi.org/10.1075/lald.53.05kea

King, J., Andrews, C. & Wagers, M. (2012). "Do reflexives always find a grammatical antecedent for themselves?" *Poster Presented at the 25th CUNY Conference on Human Sentence Processing.* New York, US.

Konnelly, L., & Cowper, E. (2020). Gender diversity and morphosyntax: An account of singular they. *Glossa: A Journal of General Linguistics*, *5*(1),1–19. https://doi.org/10.5334/GJGL.1000

Kratzer, A. (2009). Making a Pronoun: Fake Indexicals as Windows into the Properties of Pronouns. *Linguistic Inquiry*, *40*(2), 187–237. https://doi.org/10.1162/LING.2009.40.2.187

Kush, D. (2013). *Respecting relations: memory access and antecedent retrieval* (PhD Thesis, University of Maryland). Retrieved from http://www.colinphillips.net/wp-content/uploads/2014/08/kush2013.pdf

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/JSS.V082.I13

Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement attraction in Spanish comprehension. *Journal of Memory and Language*, *82*, 133–149. https://doi.org/10.1016/j.jml.2015.02.002

Lee, E.-K. R., & Phillips, C. (2022). Why non-native speakers sometimes outperform native speakers in agreement processing. *Bilingualism: Language and Cognition*, *26*(1), 152–164. https://doi.org/10.1017/s1366728922000414

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native Language Influences on Word Recognition in a Second Language: A Megastudy. *Journal of Experimental Psychology*, *34*(1), 12–31. https://doi.org/10.1037/0278-7393.34.1.12

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454. https://doi.org/10.1016/j.tics.2006.08.007

Lewis, S., & Phillips, C. (2014). Aligning Grammatical Theories and Language Processing Models. *Journal of Psycholinguistic Research*, *44*(1), 27–46. https://doi.org/10.1007/s10936-014-9329-z

Lim, J. H., & Christianson, K. (2014). Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation. *Applied Psycholinguistics*, *36*(6), 1283–1315. https://doi.org/10.1017/S0142716414000290

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, *4*(2), 126–141. https://doi.org/10.1093/APPLIN/4.2.126

Ma, Z., Wu, S., & Xu, S. (2022). Acceptance and Online Interpretation of "Gender-Neutral

Pronouns": Performance Asymmetry by Chinese English as a Foreign Language Learners. *Frontiers in Psychology*, *12*(Article 765777), 1–14. https://doi.org/10.3389/FPSYG.2021.765777/BIBTEX

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, *27*(1), 53–78. https://doi.org/10.1017/S0272263105050035

Marinis, Theodoros, & Cunnings, I. (2018). Using Psycholinguistic Techniques in a Second Language Teaching Setting. In C. Wright, T. Piske, & M. Young-Scholten (Eds.), *Mind Matters in SLA* (pp. 185–202). Multilingual Matters: Bristol. https://doi.org/10.21832/9781788921626-012/HTML

Martin, A. E., & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, *58*(3), 879–906. https://doi.org/10.1016/j.jml.2007.06.010

Martin, A. E., & McElree, B. (2009). Memory Operations That Support Language Comprehension: Evidence From Verb-Phrase Ellipsis. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*(5), 1231–1239. https://doi.org/10.1037/a0016271

Martin, A. E., & McElree, B. (2011). Direct-access retrieval during sentence comprehension: Evidence from Sluicing. *Journal of Memory and Language*, *64*(4), 327–343. https://doi.org/10.1016/j.jml.2010.12.006

McCarthy, C. (2008). Morphological variability in the comprehension of agreement: an argument for representation over computation. *Second Language Research*, *24*(4), 459–486. https://doi.org/10.1177/0267658308095737

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*(3), 381–401. https://doi.org/10.1016/j.jml.2006.06.006

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, *29*(2), 111–123. https://doi.org/10.1023/A:1005184709695

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*(1), 67–91. https://doi.org/10.1016/S0749-596X(02)00515-6

Mohammad, M. (2000). *Word Order, Agreement and Pronomalization in Standard and Palestinian Arabic.* Amsterdam:Philadelphia.: John Benjamins Publishing Company.

Mohammad, M. A. (1989). *The sentential structure of Arabic (*PhD Thesis, University of Southern California).

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, *58*(3), 503–553. https://doi.org/10.1111/j.1467-9922.2008.00449.x

Musabhien, M. (2008). *Case, Agreement and Movement in Arabic: A Minimalist Approach* (PhD Thesis, Newcastle University). Retrieved from https://theses.ncl.ac.uk/jspui/handle/10443/2046

Nakano, Y., & Wang, M. (2011). Relative-Clause Attachments in L1 and L2 English of Japanese Learners: An Offline Questionnaire and Eye-Tracking Studies. *Poster Presented at Architectures and Mechanisms for Language Processing (AMLaP)*. York, UK.

Nicenboim, B., & Vasishth, S. (2018). Models of retrieval in sentence comprehension: A computational evaluation using Bayesian hierarchical modeling. *Journal of Memory and Language, 99*, 1–34. https://doi.org/10.1016/j.jml.2017.08.004

Nicenboim, B., Vasishth, S., Engelmann, F., & Suckow, K. (2018). Exploratory and Confirmatory Analyses in Sentence Processing: A Case Study of Number Interference in German. *Cognitive Science, 42*, 1075–1100. https://doi.org/10.1111/cogs.12589

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience, 17*(8), 1212–1228. https://doi.org/10.1162/0898929055002436

Omaki, A. (2005). *Working memory and relative clause attachment in first and second language processing* (MA Thesis, University of Hawaii). Retrieved from https://scholarspace.manoa.hawaii.edu/items/e7128b0f-7878-456c-9f6b-1e11513ed33b

Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition, 33*(4), 563–588. https://doi.org/10.1017/S0272263111000313

Osterhout, L., & Mobley, L. A. (1995). Event-Related Brain Potentials Elicited by Failure to Agree. *Journal of Memory and Language, 34*(6), 739–773. https://doi.org/10.1006/JMLA.1995.1033

Parker, D., Shvartsman, M., & Van Dyke, J. (2017). The cue-based retrieval theory of sentence comprehension: new findings new challenges. In L. Escobar, V. Torrens, & T. Parodi (Eds.), *Language processing and disorders* (pp. 121–144). Newcastle: Cambridge Scholars Publishing.

Parker, D. & Phillips, C. (2014). "Selective priority for structure in memory retrieval." *Poster*

Presented at the 27th CUNY Conference on Human Sentence Processing. Columbus, OH.

Parker, D., & An, A. (2018). Not all phrases are equally attractive: Experimental evidence for selective agreement attraction effects. *Frontiers in Psychology, 9*(Article 1566), 1–16. https://doi.org/10.3389/fpsyg.2018.01566

Parker, D., & Phillips, C. (2016). Negative polarity illusions and the format of hierarchical encodings in memory. *Cognition, 157*, 321–339. https://doi.org/10.1016/j.cognition.2016.08.016

Parker, D., & Phillips, C. (2017). Reflexive attraction in comprehension is selective. *Journal of Memory and Language, 94*, 272–290. https://doi.org/10.1016/j.jml.2017.01.002

Patil, U., Vasishth, S., & Lewis, R. L. (2016). Retrieval interference in syntactic processing: The case of reflexive binding in English. *Frontiers in Psychology, 7*(Article 329), 1–18. https://doi.org/10.3389/fpsyg.2016.00329

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement Processes in Sentence Comprehension. *Journal of Memory and Language*, *41*(3), 427–456. https://doi.org/10.1006/jmla.1999.2653

Phillips, C., Wagers, M. W., & Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In J. T. Runner (Ed.), *Experiments at the Interfaces Syntax and Semantics* (pp. 147–180). Leiden: Boston: BRILL.

Pliatsikas, C., & Marinis, T. (2013). Processing empty categories in a second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition*, *16*(1), 167–182. https://doi.org/10.1017/S136672891200017X

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the*

English Language. New York: Longman.

R Development Core Team. (2019). *R: A Language and Environment for Statistical Computing.* Retrieved from https://www.r-project.org/

Radford, A. (2004). *English syntax: An introduction.*Cambridge, UK: Cambridge University Press.

Reinhart, T. (1983). *Anaphora and semantic interpretation.* London, UK: Routledge.

Ritter, E. (1993). Where's Gender? *Linguistic Inquiry*, *24*(4), 795–803. http://www.jstor.org/stable/4178843

Roberts, L. (2013). Sentence processing in Bilinguals. In Roger P. G. van Gompel (Ed.), *Sentence Processing* (pp. 221–246). London, UK: Psychology Press. https://doi.org/10.4324/9780203488454

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *18*(12), 2030–2048. https://doi.org/10.1162/jocn.2006.18.12.2030

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, *120*(8), 2022–2039. https://doi.org/10.1016/j.lingua.2010.02.004

Sagarra, N., & Herschensohn, J. (2013). Processing of gender and number agreement in late Spanish bilinguals. *International Journal of Bilingualism*, *17*(5), 607–627. https://doi.org/10.1177/1367006912453810

Schlueter, Z., Williams, A., & Lau, E. (2018). Exploring the abstractness of number retrieval

cues in the computation of subject-verb agreement in comprehension. *Journal of Memory and Language, 99*, 74–89. https://doi.org/10.1016/j.jml.2017.10.002

Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*(3), 369–385. https://doi.org/10.1017/S0142716400010845

Speyer, L. G., & Schleef, E. (2019). Processing 'Gender-neutral' Pronouns: A Self-paced Reading Study of Learners of English. *Applied Linguistics*, *40*(5), 793–815. https://doi.org/10.1093/APPLIN/AMY022

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. *Second Language Research*, *25*(1), 13–41. https://doi.org/10.1177/0267658308098995

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*(3), 542–562. https://doi.org/10.1016/S0749-596X(02)00536-3

Tahir, I. (2009). Copula in Standard English and its Counterpart in Standard Arabic. *Al-Faith Journal*, *5*(39), 74–88. https://www.iasj.net/iasj/download/1c7109937b93446e

Tanner, D., Nicol, J., & Brehm, L. (2014). The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language, 76*, 195–215. https://doi.org/10.1016/j.jml.2014.07.003

Tanner, D., Nicol, J., Herschensohn, J., & Osterhout, L. (2012). Electrophysiological markers of interference and structural facilitation in native and nonnative agreement processing. In A. Biller, E. Chung, & A. Kimball (Eds.), *Proceedings of the 36th annual Boston University*

Conference on Language Development (pp. 594–606). Somerville, MA: Cascadilla Press.

Tsimpli, I., & Dimitrakopoulou, M. (2007). The Interpretability Hypothesis: evidence from "wh"-interrogatives in second language acquisition. *Second Language Research*, *23*(2), 215–242. https://doi.org/10.1177/0267658307076546

Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing number in the real-time processing of agreement: Self-paced reading evidence from Arabic. *Frontiers in Psychology*, *6*(Article 347), 1–21. https://doi.org/10.3389/fpsyg.2015.00347

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(2), 407–430. https://doi.org/10.1037/0278-7393.33.2.407

Van Dyke, J. A., Johns, C. L., & Kukona, A. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*, *131*(3), 373–403. https://doi.org/10.1016/j.cognition.2014.01.007

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*(3), 285–316. https://doi.org/10.1016/S0749-596X(03)00081-0

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*(2), 157–166. https://doi.org/10.1016/j.jml.2006.03.007

Van Dyke, J. A., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, *65*(3), 247–263. https://doi.org/10.1016/j.jml.2011.05.002

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50*(3), 289–335. https://doi.org/10.1016/j.jml.2003.10.003

Vasishth, S., Brussow, S., Lewis, R., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science, 32*(4), 685–712. https://doi.org/10.1080/03640210802066865

Vasishth, S., & Gelman, A. (2021). How to embrace variation and accept uncertainty in linguistic and psycholinguistic data analysis. *Linguistics*, *59*(5), 1311–1342. https://doi.org/10.1515/ling-2019-0051

Vasishth, S., Nicenboim, B., Engelmann, F., & Burchert, F. (2019). Computational Models of Retrieval Processes in Sentence Processing. *Trends in Cognitive Sciences*, *23*(11), 968–982. https://doi.org/10.1016/j.tics.2019.09.003

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension : Representations and processes. *Journal of Memory and Language*, *61*(2), 206–237. https://doi.org/10.1016/j.jml.2009.04.002

Wagers, M. W., & Phillips, C. (2014). Going the distance: Memory and control processes in active dependency construction. *Quarterly Journal of Experimental Psychology*, *67*(7), 1274–1304. https://doi.org/10.1080/17470218.2013.858363

White, L. (2003). *Second Language Acquisition and Universal Grammar (Cambridge Textbooks in Linguistics)*. New York: Cambridge University Press.

White, L., Valenzuela, E., Kozlowska-Macgregor, M., & Leung, Y. K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, *25*(1), 105–133. https://doi.org/10.1017/S0142716404001067

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, *108*(1), 40–55. https://doi.org/10.1016/j.bandl.2008.10.002

Yadav, H., Smith, G., Reich, S., & Vasishth, S. (2023). Number feature distortion modulates cue-based retrieval in reading. *Journal of Memory and Language*, *129*, 1–40. https://doi.org/10.1016/j.jml.2022.104400

Zehr, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments (IBEX)*. https://doi.org/10.17605/OSF.IO/MD832

Zhou, H., Rossi, S., & Chen, B. (2017). Effects of working memory capacity and tasks in processing L2 complex sentence: Evidence from Chinese-english bilinguals. *Frontiers in Psychology*, *8*(Article 595), 1–10. https://doi.org/10.3389/fpsyg.2017.00595

Appendices

Appendix A

Background Questionnaire for L1 English Speakers

- E-mail address: …….....
- Age: ……
- Gender: Male ☐ Female ☐
- Nationality: ……
- Native language(s) (If you have learnt more than one language since birth, please list all of them): ……..…….
- Do you speak any foreign language(s)? If yes, please list all of them.

  ☐ Yes (……) No ☐
- Do you consider yourself bilingual? (i.e., Do you feel that you have a native-like command of two or more languages?). If yes, which languages?

  ☐ Yes (……) No ☐
- Were you born in the UK? ☐ Yes ☐ No
  If no, where were you born? …….
- Where did you grow up? ……
- What is your study right now?

  ☐ Bachelor degree ☐ Master degree ☐ PhD ☐ Other

Background Questionnaire for L1 Arabic Speakers

- E-mail address: ……..…..
- Age: ……
- Gender: Male ☐ Female ☐
- Nationality: ……
- Native language(s) (If you have learnt more than one language since birth, please list all of them): ……..…….
- Do you speak any foreign language(s)? If yes, please list all of them.

  ☐ Yes (……) No ☐
- Do you consider yourself bilingual? (i.e., Do you feel that you have a native-like command of two or more languages?). If yes, which languages?

  ☐ Yes (……) No ☐
- Were you born in the UK? ☐ Yes ☐ No
  If no, where were you born? …….
- When did you move to the UK? …… (month / year)
- How old were you when you started learning English? ……
- How many years or months have you lived in England or any other English-speaking country? ……
- Please tick any of the following that apply to you:

  ☐ I had English lessons at school in my home country.
  ☐ I had English lessons at university in my home country.
  ☐ I have attended English courses in an English-speaking country.
  ☐ I lived in an English-speaking country as a child
  (From age …….. until age…..……)
- Have you taken IELTS or TOEFL? ☐ Yes ☐ No

If yes**,** what is your last score? .....

When was it? .....

- What language do you speak in your house and with your family?

  …….....

- How many hours approximately per day do you spend using English outside the classroom/work environment?

  ……

- What is your study right now?

  ☐English Course ☐Bachelor degree ☐Master degree ☐PhD ☐Other

Appendix B

Agreement

1 a. The boys were hurt yesterday afternoon.

1 b. *The boy were hurt yesterday afternoon.

1 c. The boys near the girl were hurt yesterday afternoon.

1 d. *The boy near the girls were hurt yesterday afternoon.

2 a. The coaches were upset during the last match.

2 b. *The coach were upset during the last match.

2 c. The coaches of the player were upset during the last match.

2 d. *The coach of the players were upset during the last match.

3 a. The kids were delighted last weekend.

3 b. *The kid were delighted last weekend.

3 c. The kids with the clown were delighted last weekend.

3 d. *The kid with the clowns were delighted last weekend.

4 a. The kids were happy during the holidays.

4 b. *The kid were happy during the holidays.

4 c. The kids by the lady were happy during the holidays.

3 d. *The kid by the ladies were happy during the holidays.

5 a. The teachers were correct this morning.

5 b. *The teacher were correct this morning.

5 c. The teachers of the student were correct this morning.

5 d. *The teacher of the students were correct this morning.

6 a. The grandfathers were applauded last night.

6 b. *The grandfather were applauded last night.

6 c. The grandfathers of the girl were applauded last night.

6 d. *The grandfather of the girls were applauded last night.

7 a. The actors were practising two hours ago.

7 b. *The actor were practising two hours ago.

7 c. The actors behind the director were practising two hours ago.

7 d. *The actor behind the directors were practising two hours ago.

8 a. The authors were critical in the last report.

8 b. *The author were critical in the last report.

8 c. The authors by the editor were critical in the last report.

8 d. *The author by the editors were critical in the last report.

9 a. The boys were neglected yesterday evening.

9 b. *The boy were neglected yesterday evening.

9 c. The boys behind the babysitter were neglected yesterday evening.

9 d. *The boy behind the babysitters were neglected yesterday evening.

10 a. The politicians were introduced during the last meeting.

10 b. *The politician were introduced during the last meeting.

10 c. The politicians near the reporter were introduced during the last meeting.

10 d. *The politician near the reporters were introduced during the last meeting.

11 a. The girls were rescued from a kidnapper yesterday.

11 b. *The girl were rescued from a kidnapper yesterday.

11 c. The girls with the cop were rescued from a kidnapper yesterday.

11 d. * The girl with the cops were rescued from a kidnapper yesterday.

12 a. The schoolboys were respectful in the last class.

12 b. *The schoolboy were respectful in the last class.

12 c. The schoolboys behind the teacher were respectful in the last class.

12 d. *The schoolboy behind the teachers were respectful in the last class.

13 a. The cops were protective last night.

13 b. *The cop were protective last night.

13 c. The cops by the neighbour were protective last night.

13 d. *The cop by the neighbours were protective last night.

14 a. The guards were identified as suspicious yesterday.

14 b. *The guard were identified as suspicious yesterday.

14 c. The guards near the judge were identified as suspicious yesterday.

14 d. *The guard near the judges were identified as suspicious yesterday.

15 a. The sisters were hurt during Saturday's incident.

15 b. *The sister were hurt during Saturday's incident.

15 c. The sisters of the boy were hurt during Saturday's incident.

15 d. *The sister of the boys were hurt during Saturday's incident.

16 a. The grandmothers were trained in knitting.

16 b. *The grandmother were trained in knitting.

16 c. The grandmothers of the boy were trained in knitting.

16 d. *The grandmother of the boys were trained in knitting.

17 a. The riders were injured in the race yesterday.

17 b. *The rider were injured in the race yesterday.

17 c. The riders behind the princess were injured in the race yesterday.

17 d. *The rider behind the princesses were injured in the race yesterday.

18 a. The widows were very helpful.

18 b. *The widow were very helpful .

18 c. The widows near the customer were helpful.

18 d. *The widow near the customers were very helpful.

19 a. The daughters were clothed after swimming in the afternoon.

19 b. *The daughter were clothed after swimming in the afternoon.

19 c. The daughters of the parent were clothed after swimming in the afternoon.

19 d. *The daughter of the parents were clothed after swimming in the afternoon.

20 a. The brides were promised a happy life last night.

20 b. *The bride were promised a happy life last night.

20 c. The brides behind the waiter were promised a happy life last night.

20 d. *The bride behind the waiters were promised a happy life last night.

21 a. The maids were surprised by the gifts last night.

21 b. *The maid were surprised by the gifts last night.

21 c. The maids of the prince were surprised by the gifts last night.

21 d. *The maid of the princes were surprised by the gifts last night.

22 a. The schoolboys were awarded a medal last week.

22 b. *The schoolboy were awarded a medal last week.

22 c. The schoolboys behind the schoolgirl were awarded a medal last week.

22 d. *The schoolboy behind the schoolgirls were awarded a medal last week.

23 a. The workers were wounded after the accident yesterday.

23 b. *The worker were wounded after the accident yesterday.

23 c. The workers near the maid were wounded after the accident yesterday.

23 d. *The worker near the maids were wounded after the accident yesterday.

24 a. The daughters were lovely in the wedding last night.

24 b. *The daughter were lovely in the wedding last night.

24 c. The daughters near the boy were lovely in the wedding last night.

24 d. *The daughter near the boys were lovely in the wedding last night.

Reflexive

1 a. The maid cleaned herself up after finishing the gardening.

1 b. *The maid cleaned himself up after finishing the gardening.

1 c. The maid with the butler cleaned herself up after finishing the gardening.

1 d. *The maid with the butler cleaned himself up after finishing the gardening.

2 a. The sister covered herself with a blanket last night.

2 b. *The sister covered himself with a blanket last night.

2 c. The sister of the boy covered herself with a blanket last night.

2 d. *The sister of the boy covered himself with a blanket last night.

3 a. The businesswoman satisfied herself with the sales growth figures last month.

3 b. *The businesswoman satisfied himself with the sales growth figures last month.

3 c. The businesswoman behind the salesman satisfied herself with the sales growth figures last month.

3 d. *The businesswoman behind the salesman satisfied himself with the sales growth figures last month.

3 a. The congresswoman embarrassed herself yesterday afternoon.

4 b. *The congresswoman embarrassed himself yesterday afternoon.

4 c. The congresswoman near the waiter embarrassed herself yesterday afternoon.

4 d. *The congresswoman near the waiter embarrassed himself yesterday afternoon.

5 a. The policeman recognized himself in the picture yesterday.

5 b. *The policeman recognized herself in the picture yesterday.

5 c. The policeman with the lady recognized himself in the picture yesterday.

4 d. *The policeman with the lady recognized herself in the picture yesterday.

6 a. The boy entertained himself at the last party.

6 b. *The boy entertained herself at the last party.

6 c. The boy near the lady entertained himself at the last party.

6 d. *The boy near the lady entertained herself at the last party.

7 a. The schoolboy found himself lost yesterday morning.

7 b. *The schoolboy found herself lost yesterday morning.

7 c. The schoolboy behind the policewoman found himself lost yesterday morning.

7 d. *The schoolboy behind the policewoman found herself lost yesterday morning.

8 a. The prince surrounded himself with guards yesterday.

8 b. *The prince surrounded herself with guards yesterday.

8 c. The prince by the queen surrounded himself with guards yesterday.

8 d. *The prince by the queen surrounded herself with guards yesterday.

9 a. The grandmother entertained herself with a catalogue of last year's fashion.

9 b. *The grandmother entertained himself with a catalogue of last year's fashion.

9 c. The grandmother of the boy entertained herself with a catalogue of last year's fashion.

9 d. * The grandmother of the boy entertained himself with a catalogue of last year's fashion.

10 a. The businesswoman committed herself to finalizing the report yesterday.

10 b. * The businesswoman committed himself to finalizing the report yesterday.

10 c. The businesswoman near the salesman committed herself to finalizing the report yesterday.

10 d. The businesswoman near the salesman committed himself to finalizing the report yesterday.

11 a. The businesswoman supported herself during the changes.

11 b. *The businesswoman supported himself during the changes.

11 c. The businesswoman near the salesman supported herself during the changes.

11 d. *The businesswoman near the salesman supported himself during the changes.

12 a. The spokeswoman presented herself as an honest person at the event.

12 b. *The spokeswoman presented himself as an honest person at the event.

12 c. The spokeswoman for the king presented herself as an honest person at the event.

12 d. *The spokeswoman for the king presented himself as an honest person at the event.

13 a. The businessman praised himself during Sunday's celebration.

13 b. *The businessman praised herself during Sunday's celebration.

13 c. The businessman behind the waitress praised himself during Sunday's celebration.

13 d. *The businessman behind the waitress praised herself during Sunday's celebration.

14 a. The councilman repeated himself in the last meeting.

14 b. *The councilman repeated herself in the last meeting.

14 c. The councilman with the lady repeated himself in the last meeting.

14 d. *The councilman with the lady repeated herself in the last meeting.

15 a. The councilman prepared himself for the interview.

15 b. *The councilman prepared herself for the interview.

15 c. The councilman with the spokeswoman prepared himself for the interview.

15 d. * The councilman with the spokeswoman prepared herself for the interview.

16 a. The schoolboy convinced himself that the teacher was always right.

16 b. *The schoolboy convinced herself that the teacher was always right.

16 c. The schoolboy with the schoolgirl convinced himself that the teacher was always right.

16 d. *The schoolboy with the schoolgirl convinced herself that the teacher was always right.

17 a. The businesswoman encouraged herself to start a new project.

17 b. *The businesswoman encouraged himself to start a new project.

17 c. The businesswoman by the waiter encouraged herself to start a new project.

17 d. *The businesswoman by the waiter encouraged himself to start a new project.

18 a. The waitress defended herself yesterday morning.

18 b. *The waitress defended himself yesterday morning

18 c. The waitress near the chairman defended herself yesterday morning.

18 d. *The waitress near the chairman defended himself yesterday morning.

19 a. The schoolgirl amused herself during the games yesterday.

19 b. *The schoolgirl amused himself during the games yesterday.

19 c. The schoolgirl by the schoolboy amused herself during the games yesterday.

19 d. *The schoolgirl by the schoolboy amused himself during the games yesterday.

20 a. The actress hated herself in the television series.

20 b. *The actress hated himself in the television series.

20 c. The actress by the cameraman hated herself in the television series.

20 d. *The actress by the cameraman hated himself in the television series.

21 a. The congressman assigned himself a new task last year.

21 b. *The congressman assigned herself a new task last year.

21 c. The congressman behind the policewoman assigned himself a new task last year.

21 d. *The congressman behind the policewoman assigned herself a new task last year.

22 a. The workman recommended himself to the company yesterday.

22 b. *The workman recommended herself to the company yesterday.

22 c. The workman near the spokeswoman recommended himself to the company yesterday.

22 d. *The workman near the spokeswoman recommended herself to the company yesterday.

23 a. The man cut himself two hours ago.

23 b. *The man cut herself two hours ago.

23 c. The man behind the lady cut himself two hours ago.

23 d. *The man behind the lady cut herself two hours ago

24 a. The father prepared himself for the wedding ceremony last night.

24 b.*The father prepared herself for the wedding ceremony last night.

24 c. The father of the bride prepared himself for the wedding ceremony last night.

24 d. *The father of the bride prepared herself for the wedding ceremony last night.

Fillers

1. The friend of the girl played football last night.
2. The woman near the waiter usually comes in for lunch at 12.00 p.m.
3. The expert opposite the students visited our university last month.
4. Before the meeting, the secretary reminded herself of the important points.
5. When Sara finished the project yesterday, she went shopping and spoiled herself.
6. While the mother was cleaning the house, the boy injured himself with a sharp knife.
7. The children are unable to look after themselves.
8. The students said that they will organize themselves before the event.
9. Laura eats balanced meals every day to keep fit.
10. Whenever Bob is free, he reads articles on the internet.
11. My mother is preparing lunch in the kitchen now.
12. Louise was hungry by 11 o'clock, so she ate a sandwich.
13. I visited my grandmother two days ago.
14. Julia told me that she cannot attend my sister's wedding next week.
15. Sally is very quiet in class whereas her sister is very talkative.
16. My house is the largest one in our neighbourhood.
17. The businessman wants to start a new project soon.
18. I always drink two cups of coffee every day.
19. Sara does not have many friends at school.
20. We saw many animals in the zoo yesterday.
21. The students have little interest in politics.
22. The headmaster received some complaints from the students last week.
23. My brother has a dog while my sister has a cat.
24. We met some friends for drinks after work yesterday.
25. Sam took lots of photos of the mountains, but Emma did not take any.
26. I will read the article and summarize it next week.

27. Isabella and Rose are going to visit the museum next weekend.
28. The student wants to do her assignment before bedtime.
29. We play football with our neighbours once a month.
30. Sara and Rose are friends, and they have known each other for 5 years.
31. After I prepared all the ingredients, I cooked the dinner.
32. Clare doesn't like to have any cream with her coffee.
33. * The girl next to the boy will writing an essay in the library.
34. * The man behind the lady has went to the gym every morning.
35. * The engineer with the secretaries can works with aircrafts.
36. * After the dinner last night, the girl washed himself with soap.
37. * When Dan presented yesterday morning, he introduced herself as a freelance translator.
38. * While the father was watching TV, the girl taught himself math.
39. * The musician realised that Susan cut themselves.
40. * The neighbour claimed that Tom sacrificed themselves.
41. * In 2001, I have wrote my first book and it was a best seller in that year.
42. * Before I go to sleep, I watching TV for an hour every night.
43. * When my cousins visited me last weekend, I was show her my new bedroom.
44. * Rose remembered that John surprised they with a gift last month.
45. * I helped themselves to the birthday cake.
46. * Tom said that he will studies for the exam next week.
47. * Dan's cat is more large than my dog.
48. * My brother plays tennis gooder than I do.
49. * John is want a snack, but he could not find anything to eat.
50. * I had chicken and rices for lunch yesterday when I went to a near Indian restaurant.
51. * Philip owns much houses in France.
52. * The recipe does not need many milk.
53. * There were very little students studied Spanish last year.
54. * Isabella saves a few money every month.
55. * The teacher spending a hour teaching Sara calculus.

56. * I am no see some friends there on Thursday.

57. * I don't have some information about flights to Paris.

58. * I will eat fish for dinner and drank milk.

59. * Anna and Mike is go skiing tomorrow.

60. * Dan likes fish and I prefering chicken.

61. * We usually will enjoying horror movies in the cinema.

62. * Dan and Isabella are married, and he has been with each other for 20 years.

63. * I fed all of fish, then cleaned tank.

64. * I don’t wanting not pudding on the cake.

Appendix C

Agreement

1 a. The waitresses near the schoolgirls unsurprisingly were unhappy about all the noise.
1 b. The waitresses near the schoolgirl unsurprisingly were unhappy about all the noise.
1 c. *The waitress near the schoolgirls unsurprisingly were unhappy about all the noise.
1 d. *The waitress near the schoolgirl unsurprisingly were unhappy about all the noise.
Q: Was it noisy? Yes

2 a. The hosts near the guests certainly were impressed with the Christmas decorations.
2 b. The hosts near the guest certainly were impressed with the Christmas decorations.
2 c. *The host near the guests certainly were impressed with the Christmas decorations.
2 d. *The host near the guest certainly were impressed with the Christmas decorations.
Q: Were the decorations for Easter? No

3 a. The assistants behind the passengers certainly were pleased with the long flight.
3 b. The assistants behind the passenger certainly were pleased with the long flight.
3 c. *The assistant behind the passengers certainly were pleased with the long flight.
3 d. *The assistant behind the passenger certainly were pleased with the long flight.
Q: Was it a long flight? Yes

4 a. The professors of the students likely were concerned about the low grades.
4 b. The professors of the student likely were concerned about the low grades.
4 c. *The professor of the students likely were concerned about the low grades.
4 d. *The professor of the student likely were concerned about the low grades.
Q: Were the grades high? No

5 a. The inspectors by the neighbours recently were mentioned in the local newspaper.
5 b. The inspectors by the neighbour recently were mentioned in the local newspaper.
5 c. *The inspector by the neighbours recently were mentioned in the local newspaper.
5 d. *The inspector by the neighbour recently were mentioned in the local newspaper.
Q: Was/were the inspector/inspectors by the neighbours(s)? Yes

6 a. The professors near the invigilators certainly were mistaken about the final question.
6 b. The professors near the invigilator certainly were mistaken about the final question.
6 c. *The professor near the invigilators certainly were mistaken about the final question.
6 d. *The professor near the invigilator certainly were mistaken about the final question.
Q: Was/were the professor(s) correct about the final question? No

7 a. The clowns behind the kids definitely were annoyed by all the noise.
7 b. The clowns behind the kid definitely were annoyed by all the noise.
7 c. *The clown behind the kids definitely were annoyed by all the noise.
7 d. *The clown behind the kid definitely were annoyed by all the noise.

Q: Was/were the clown(s) behind the kid(s)? Yes

8 a. The journalists opposite the criminals apparently were investigated for some fraudulent claims.
8 b. The journalists opposite the criminal apparently were investigated for some fraudulent claims.
8 c. *The journalist opposite the criminals apparently were investigated for some fraudulent claims.
8 d. *The journalist opposite the criminal apparently were investigated for some fraudulent claims.
Q: Was/were the journalist(s) trustworthy? No

9 a. The directors of the assistants suddenly were intrigued by the secret report.
9 b. The directors of the assistant suddenly were intrigued by the secret report.
9 c. *The director of the assistants suddenly were intrigued by the secret report.
9 d. *The director of the assistant suddenly were intrigued by the secret report.
Q: Was the report secret? Yes

10 a. The nurses of the patients clearly were nervous about the stressful situation.
10 b. The nurses of the patient clearly were nervous about the stressful situation.
10 c. *The nurse of the patients clearly were nervous about the stressful situation.
10 d. *The nurse of the patient clearly were nervous about the stressful situation.
Q: Was it a relaxing situation? No

11 a. The teenagers by the parents clearly were happy in the sunny weather.
11 b. The teenagers by the parent clearly were happy in the sunny weather.
11 c. *The teenager by the parents clearly were happy in the sunny weather.
11 d. *The teenager by the parent clearly were happy in the sunny weather.
Q: Was it sunny? Yes

12 a. The assistants of the pilots definitely were scared about the approaching storm.
12 b. The assistants of the pilot definitely were scared about the approaching storm.
12 c. *The assistant of the pilots definitely were scared about the approaching storm.
12 d. *The assistant of the pilot definitely were scared about the approaching storm.
Q: Was the weather going to be calm? No

13 a. The boys by the girls suddenly were scared about the loud noises.
13 b. The boys by the girl suddenly were scared about the loud noises.
13 c. *The boy by the girls suddenly were scared about the loud noises.
13 d. *The boy by the girl suddenly were scared about the loud noises.
Q: Were the noises loud? Yes

14 a. The athletes behind the coaches unsurprisingly were upset about losing the championship.
14 b. The athletes behind the coach unsurprisingly were upset about losing the championship.

14 c. *The athlete behind the coaches unsurprisingly were upset about losing the championship.
14 d. *The athlete behind the coach unsurprisingly were upset about losing the championship.
Q: Was/were the athlete(s) happy about the loss? No

15 a. The clowns by the schoolboys apparently were good at drawing complex shapes.
15 b. The clowns by the schoolboy apparently were good at drawing complex shapes.
15 c. *The clown by the schoolboys apparently were good at drawing complex shapes.
15 d. *The clown by the schoolboy apparently were good at drawing complex shapes.
Q: Was/were the clown(s) by the schoolboy(s)? Yes

16 a. The soldiers opposite the terrorists likely were surrounded by two insurgent groups.
16 b. The soldiers opposite the terrorist likely were surrounded by two insurgent groups.
16 c. *The soldier opposite the terrorists likely were surrounded by two insurgent groups.
16 d. *The soldier opposite the terrorist likely were surrounded by two insurgent groups.
Q: Were there five groups of insurgents? No

17 a. The researchers near the participants clearly were happy with the experiment's results.
17 b. The researchers near the participant clearly were happy with the experiment's results.
17 c. *The researcher near the participants clearly were happy with the experiment's results.
17 d. *The researcher near the participant clearly were happy with the experiment's results.
Q: Was/were the researcher near the participant(s)? Yes

18 a. The reporters opposite the candidates definitely were nervous about the upcoming debate.
18 b. The reporters opposite the candidate definitely were nervous about the upcoming debate.
18 c. *The reporter opposite the candidates definitely were nervous about the upcoming debate.
18 d. *The reporter opposite the candidate definitely were nervous about the upcoming debate.
Q: Was/were the reporter(s) calm about the upcoming debate? No

19 a. The musicians behind the fans certainly were upset about the horrible performance.
19 b. The musicians behind the fan certainly were upset about the horrible performance.
19 c. *The musician behind the fans certainly were upset about the horrible performance.
19 d. *The musician behind the fan certainly were upset about the horrible performance.
Q: Was the performance horrible? Yes

20 a. The gardeners by the maids apparently were unable to buy more plants.
20 b. The gardeners by the maid apparently were unable to buy more plants.
20 c. *The gardener by the maids apparently were unable to buy more plants.
20 d. *The gardener by the maid apparently were unable to buy more plants.
Q: Did the gardener(s) buy more plants? No

21 a. The officers behind the firefighters allegedly were pulled over for drunk driving.
21 b. The officers behind the firefighter allegedly were pulled over for drunk driving.
21 c. *The officer behind the firefighters allegedly were pulled over for drunk driving.
21 d. *The officer behind the firefighter allegedly were pulled over for drunk driving.

Q: Was/were the officer(s) behind the firefighter(s)? Yes

22 a. The athletes near the referees obviously were disappointed about the bad call.
22 b. The athletes near the referee obviously were disappointed about the bad call.
22 c. *The athlete near the referees obviously were disappointed about the bad call.
22 d. *The athlete near the referee obviously were disappointed about the bad call.
Q: Was it a good call? No

23 a. The ministers near the teachers obviously were tempted to discuss controversial issues.
23 b. The ministers near the teacher obviously were tempted to discuss controversial issues.
23 c. *The minister near the teachers obviously were tempted to discuss controversial issues.
23 d. *The minister near the teacher obviously were tempted to discuss controversial issues.
Q: Was/were the minister(s) near the teacher(s)? Yes

24 a. The students of the teachers obviously were upset about the low grades.
24 b. The students of the teacher obviously were upset about the low grades.
24 c. *The student of the teachers obviously were upset about the low grades.
24 d. *The student of the teacher obviously were upset about the low grades.
Q: Did everybody receive high grades? No

Reflexive

1 a. The maid of the businesswoman clearly upset herself by doing other people's work.
1 b. The maid of the businessman clearly upset herself by doing other people's work.
1 c. *The handyman of the businesswoman clearly upset herself by doing other people's work.
1 d. *The handyman of the businessman clearly upset herself by doing other people's work.
Q: Was somebody upset? Yes

2 a. The girl with the woman apparently amused herself more happily than ever before.
2 b. The girl with the man apparently amused herself more happily than ever before.
2 c. *The boy with the woman apparently amused herself more happily than ever before.
2 d. *The boy with the man apparently amused herself more happily than ever.
Q: Was somebody unhappy? No

3 a. The mother of the girl unsurprisingly tired herself with the traffic on roads.
3 b. The mother of the boy unsurprisingly tired herself with the traffic on roads.
3 c. *The father of the girl unsurprisingly tired herself with the traffic on roads.
3 d. *The father of the boy unsurprisingly tired herself with the traffic on roads.
Q: Was the traffic on the roads tiring? Yes

4 a. The niece of the queen apparently hid herself far from the town centre.
4 b. The niece of the king apparently hid herself far from the town centre.
4 c. *The nephew of the queen apparently hid herself far from the town centre.
4 d. *The nephew of the king apparently hid herself far from the town centre.

Q: Was the niece/nephew hiding in the town centre? No

5 a. The policeman by the chairman apparently volunteered himself to drive the kids home.
5 b. The policeman by the chairwoman apparently volunteered himself to drive the kids home.
5 c. *The policewoman by the chairman apparently volunteered himself to drive the kids home.
5 d. *The policewoman by the chairwoman apparently volunteered himself to drive the kids home.
Q: Did the kids need to be driven home? Yes

6 a. The man near the schoolboy suddenly worried himself about the experiment and results.
6 b. The man near the schoolgirl suddenly worried himself about the experiment and results.
6 c. *The woman near the schoolboy suddenly worried himself about the experiment and results.
6 d. *The woman near the schoolgirl suddenly worried himself about the experiment and results.
Q: Was the man/woman far away from the schoolboy/girl? No

7 a. The cameraman with the prince undoubtedly prepared himself for all the negative criticism.
7 b. The cameraman with the princess undoubtedly prepared himself for all the negative criticism.
7 c. *The camerawoman with the prince undoubtedly prepared himself for all the negative criticism.
7 d. *The camerawoman with the princess undoubtedly prepared himself for all the negative criticism.
Q: Was the criticism negative? Yes

8 a. The gentleman near the cameraman apparently excused himself from taking the big job.
8 b. The gentleman near the camerawoman apparently excused himself from taking the big job.
8 c. *The lady near the cameraman apparently excused himself from taking the big job.
8 d. *The lady near the camerawoman apparently excused himself from taking the big job.
Q: Did the gentleman/lady agree to take the big job? No

9 a. The schoolgirl with the chairwoman unsurprisingly proved herself during the big audition yesterday.
9 b. The schoolgirl with the chairman unsurprisingly proved herself during the big audition yesterday.
9 c. *The schoolboy with the chairwoman unsurprisingly proved herself during the big audition yesterday.
9 d. *The schoolboy with the chairman unsurprisingly proved herself during the big audition yesterday.
Q: Was there an audition yesterday? Yes

10 a. The aunt of the girl certainly declared herself interested in the exciting news.
10 b. The aunt of the boy certainly declared herself interested in the exciting news.
10 c. *The uncle of the girl certainly declared herself interested in the exciting news.
10 d. *The uncle of the boy certainly declared herself interested in the exciting news.
Q: Was the news boring? No

11 a. The grandmother of the girl probably supported herself during the very difficult times.
11 b. The grandmother of the boy probably supported herself during the very difficult times.
11 c. *The grandfather of the girl probably supported herself during the very difficult times.
11 d. *The grandfather of the boy probably supported herself during the very difficult times.
Q: Were times difficult? Yes

12 a. The businesswoman near the girl evidently rescued herself from accidently hitting the car.
12 b. The businesswoman near the boy evidently rescued herself from accidently hitting the car.
12 c. *The businessman near the girl evidently rescued herself from accidently hitting the car.
12 d. *The businessman near the boy evidently rescued herself from accidently hitting the car.
Q: Was somebody hit by a truck? No

13 a. The nephew of the groom completely cheered himself up during the big day.
13 b. The nephew of the bride completely cheered himself up during the big day.
13 c. *The niece of the groom completely cheered himself up during the big day.
13 d. *The niece of the bride completely cheered himself up during the big day.
Q: Did the groom/bride have a nephew/niece? Yes

14 a. The groom behind the boy excitedly reminded himself about the recent good news.
14 b. The groom behind the girl excitedly reminded himself about the recent good news.
14 c. *The bride behind the boy excitedly reminded himself about the recent good news.
14 d. *The bride behind the girl excitedly reminded himself about the recent good news.
Q: Was the recent news bad? No

15 a. The stepfather of the groom unsurprisingly prepared himself for the big day ahead.
15 b. The stepfather of the bride unsurprisingly prepared himself for the big day ahead.
15 c. *The stepmother of the groom unsurprisingly prepared himself for the big day ahead.
15 d. *The stepmother of the bride unsurprisingly prepared himself for the big day ahead.
Q: Was it going to be a big day? Yes

16 a. The prince opposite the king assertively described himself as a human rights defender.
16 b. The prince opposite the queen assertively described himself as a human rights defender.
16 c. *The princess opposite the king assertively described himself as a human rights defender.
16 d. *The princess opposite the queen assertively described himself as a human rights defender.
Q: Was the prince/princess against human rights? No

17 a. The princess near the congresswoman clearly enjoyed herself while reading the victory story.
17 b. The princess near the congressman clearly enjoyed herself while reading the victory story.
17 c. *The prince near the congresswoman clearly enjoyed herself while reading the victory story.
17 d. *The prince near the congressman clearly enjoyed herself while reading the victory story.
Q: Was the princess/prince near the congresswoman/man? Yes

18 a. The stateswoman opposite the camerawoman undoubtedly reassured herself that the event would succeed.
18 b. The stateswoman opposite the cameraman undoubtedly reassured herself that the event would succeed.
18 c. *The statesman opposite the camerawoman undoubtedly reassured herself that the event would succeed.
18 d. *The statesman opposite the cameraman undoubtedly reassured herself that the event would succeed.
Q: Was the stateswoman/man opposite the president? No

19 a. The lady behind the spokeswoman apparently prepared herself for the live musical performance.
19 b. The lady behind the spokesman apparently prepared herself for the live musical performance.
19 c. *The man behind the spokeswoman apparently prepared herself for the live musical performance.
19 d. *The man behind the spokesman apparently prepared herself for the live musical performance.
Q: Was there a musical performance? Yes

20 a. The maid by the woman apparently bought herself more than 100 seasonal flowers.
20 b. The maid by the man apparently bought herself more than 100 seasonal flowers.
20 c. *The butler by the woman apparently bought herself more than 100 seasonal flowers.
20 d. *The butler by the man apparently bought herself more than 100 seasonal flowers.
Q: Did the maid/butler buy some onions? No

21 a. The gentleman by the salesman clearly found himself charged for the wrong items.
21 b. The gentleman by the saleswoman clearly found himself charged for the wrong items.
21 c. *The lady by the salesman clearly found himself charged for the wrong items.
21 d. *The lady by the saleswoman clearly found himself charged for the wrong items.
Q: Was somebody charged for the wrong items? Yes

22 a. The businessman by the salesman taught himself some competitive sales management strategies.
22 b. The businessman by the saleswoman taught himself some competitive sales management strategies.
22 c. *The businesswoman by the salesman taught himself some competitive sales management strategies.
22 d. *The businesswoman by the saleswoman taught himself some competitive sales management strategies.
Q: Were the sales management strategies uncompetitive? No

23 a. The man near the policeman strangely isolated himself from society for many years.
23 b. The man near the policewoman strangely isolated himself from society for many years.

23 c. *The lady near the policeman strangely isolated himself from society for many years.
23 d. *The lady near the policewoman strangely isolated himself from society for many years.
Q: Was the man/lady near the policeman/policewoman? Yes

24 a. The man by the prince obviously prepared himself for the big ceremony tonight.
24 b. The man by the princess obviously prepared himself for the big ceremony tonight.
24 c. *The lady by the prince obviously prepared himself for the big ceremony tonight.
24 d. *The lady by the princess obviously prepared himself for the big ceremony tonight.
Q: Was the man/woman far from the prince/princes? No

Fillers

1. The manager of the employees unsurprisingly was upset about the late report.
Q: Were the employees upset? No

2. The woman with the coach unsurprisingly was the winner of the race.
Q: Did the woman win? Yes

3. The optometrist near the doctors clearly was happy with the patient's corrected vision.
Q: Was the optometrist near the doctors? Yes

1. The hairstylist by the barber surprisingly was even more beautiful in person.
Q: Was it the barber who was beautiful? No

2. The brothers of the girl happily walked home from the cinema.
Q: Were the brothers in the supermarket? No

6.The teachers of the students patiently waited for the correct answers.
Q: Were the teachers patient? Yes

7. The waiters near the woman gratefully accepted the tip.
Q: Did the waiters get tipped? Yes

8. The hikers with the officers clearly enjoyed the view from the mountain peak.
Q: Were the hikers at the top of a mountain? Yes

9. The prince by the queen happily presented the trophy to the winner.
Q: Was there a winner? Yes

10. The man behind the boy accidently hurt the woman during the office party.
Q: Was somebody hurt? Yes

11. The mother of the boy happily cooked the food for dinner.
Q: Was the food for lunch? No

12. The king opposite the princes criticised the minister for his actions.
Q: Did the king criticize the minister? Yes

13.The princess with the camerawoman confidently criticized the media for the fake news.
Q: Did the princess blame the media? Yes

14. After praising herself, the headmistress behaved cruelly to the people close to her.
Q: Was the headmistress nice? No

15. Before introducing herself to the audience, the lady took a deep breath.
Q: Did the lady drink water before the presentation? No

16. Before surrendering himself to the police, the killer waited in an abandoned house.
Q: Did the killer give himself up to the police? Yes

17. After exhausting herself during the day, the grandmother couldn't leave the house.
Q: Was the grandmother able to leave the house? No

18. After preparing himself for the championship game, the schoolboy promised to do his best to win.
Q: Was the schoolboy ready for the championship game? Yes

19. While the boy was cleaning himself of mud, the mother was preparing new clothes for him.
Q: Was the mother cleaning the boy? No

20. While the girl was washing herself after school, the mother was cooking the dinner.
Q: Was the mother washing the girl? No

21. George mentioned that Sara invited him to the dinner.
Q: Was George invited to the dinner? Yes

22. John said that Dan surprised him with a ticket to New York.
Q: Did Dan buy a ticket to London? No

23. Tom remembered that Rose asked him about the timing of his book release.
Q: Did Rose ask about the movie release? No

24. James thinks that Adam blames him after they both skipped yesterday's class.
Q: Did Adam attend the class yesterday? No

25. Sophia thought that George would invite her to the party.

Q: Did Sophie expect an invitation to the party? Yes

26. Isabella said that Sara reminded her to phone the company.
Q: Was Isabella reminded to call the company? Yes

27. Rose said that John mentioned her in the newspaper as an inspiring person.
Q: Was Rose seen as a boring person? No

28. Emily knows that Sara sent her an email asking for some advice.
Q: Did Emily receive an email from Sara? Yes

29. When she got home, Maria sent Tom a message.
Q: Did Maria give Tom a call? No

30. When he worked for the bank, Dan used to skip lunch.
Q: Did Dan used to skip lunch? Yes

31. Before she bought a flat, Sara asked for some recommendations.
Q: Did Sara buy a car? No

32. After he finished work, John visited his friend in the hospital.
Q: Was John's friend in the hospital? Yes

33. She said that Adam bought Tom a gift.
Q: Did Adam buy a gift for Tom? Yes

34. He thought that David lent Ali a book.
Q: Did David lend Ali a bicycle? No

35. They thought that Rose gave Emily some chocolate.
Q: Was Emily given some flowers? No

36. The teacher assured that each child received a present.
Q: Did each child receive a present? Yes

37. The editor mentioned that the journal is published every month.
Q: Is the journal published monthly? Yes

38. Isabella said that she drinks her coffee here every day.
Q: Does Isabella drink coffee daily? Yes

39. The teacher reported that the students blamed themselves for not passing the exam.
Q: Did the students fail the exam? Yes

40. The presenter said that the dancers introduced themselves.
Q: Did the presenter introduce the dancers? No

41. It was the Chemistry assignment that Sara submitted yesterday.
Q: Did Sara submit the Biology assignment yesterday? No

42. Blueberries that are freshly picked are the most delicious.
Q: Are fresh blueberries delicious? Yes

43. The researcher who spilled the chemicals quickly contacted his supervisor.
Q: Was the supervisor contacted? Yes

44. John hesitated a lot, but he shouldn't anymore.
Q: Was John hesitant? Yes

45. Kevin asked which bread Steven usually buys in the supermarket.
Q: Was Kevin looking for milk in the supermarket? No

46. I met the nurse who ignored the patient.
Q: Did the nurse help the patient? No

47. The ball was kicked by the boy in the street.
Q: Was it the boy who kicked the ball? Yes

48. Jasmine can draw an elephant, but Ryan can't.
Q: Can Ryan draw an elephant? No

49. John liked the book that the girl in the library read yesterday.
Q: Did the girl read the book in the train yesterday? No

50. John wondered why the bus is always late.
Q: Does the bus come on time? No

51. The pilot who flew the planes across the ocean had a lot of experience.
Q: Was the pilot experienced? Yes

52. The article was assessed and reviewed by the editor.
Q: Did the writer review the article? No

53. It was because Paul was ill that we decided to return.
Q: Was Paul ill? Yes

54. The fool that was arrested by the cops was running naked in the street.
Q: Was the fool clothed? No

55.The technician that the computer confused was being paid 15 dollars per hour.
Q: Was the technician payed 30 dollars per hour? No

56. It was from John that she heard the news.
Q: Did she hear the news from John? Yes

57. The girl that kissed the parents was loving and sweet to her friends.
Q: Was the girl nice? Yes

58. What the lady ordered was the cake made of chocolate.
Q: Did the lady order vanilla cake? No

59. It was the coffee that the man ordered in the café.
Q: Did the man order tea in the café? No

60. After the long race, Kevin was tired and Susan was too.
Q: Was the race long? Yes

61. John saw the officer that arrested the robber.
Q: Was the robber arrested? Yes

62. The man that designed the building was not interested in his wife anymore.
Q: Was the man interested in his wife? No

63.The author that was encouraged by the researcher was writing a science book
Q: Was it a religious book that the author was writing? No

64. Karen should fly to Paris, and her mother should too.
Q: Should Karen fly to London? No

Appendix D

1a. The boys hurt themselves yesterday afternoon.
1b. *The boy hurt themselves yesterday afternoon.
1c. The boys near the girl hurt themselves yesterday afternoon.
1d. *The boy near the girls hurt themselves yesterday afternoon.

2 a. The kids entertained themselves last weekend.
2 b. *The kid entertained themselves last weekend.
2 c. The kids with the housekeeper entertained themselves last weekend.
2 d. *The kid with the housekeepers entertained themselves last weekend.

3 a. The students taught themselves yesterday morning.
3 b. *The student taught themselves yesterday morning.
3 c. The students of the teacher taught themselves yesterday morning.
3 d. *The student of the teachers taught themselves yesterday morning.

4 a. The actors presented themselves two hours ago.
4 b. *The actor presented themselves two hours ago.
4 c. The actors behind the director presented themselves two hours ago.
4 d. *The actor behind the directors presented themselves two hours ago.

5 a. The authors criticised themselves in the last report.
5 b. *The author criticised themselves in the last report.
5 c. The authors by the editor criticised themselves in the last report.
5 d. *The author by the editors criticised themselves in the last report.

6 a. The politicians introduced themselves during the last meeting.
6 b. *The politician introduced themselves during the last meeting.
6 c. The politicians near the reporter introduced themselves during the last meeting.
6 d. *The politician near the reporters introduced themselves during the last meeting.

7 a. The girls saved themselves from a kidnapper yesterday.
7 b. *The girl saved themselves from a kidnapper yesterday.
7 c. The girls with the cop saved themselves from a kidnapper yesterday.
7 d. * The girl with the cops saved themselves from a kidnapper yesterday.

8 a. The neighbours protected themselves last night.
8 b. *The neighbour protected themselves last night.
8 c. The neighbours by the cop protected themselves last night.
8 d. *The neighbour by the cops protected themselves last night.

9 a. The guards identified themselves as police yesterday.
9 b. *The guard identified themselves as police yesterday.
9 c. The guards near the judge identified themselves as police yesterday.
9 d. *The guard near the judges identified themselves as police yesterday.

10 a. The sisters hurt themselves during gym class.
10 b. *The sister hurt themselves during gym class.
10 c. The sisters of the boy hurt themselves during gym class.
10 d. *The sister of the boys hurt themselves during gym class.

11 a. The grandmothers taught themselves to knit.
11 b. *The grandmother taught themselves to knit.
11 c. The grandmothers of the boy taught themselves to knit.
11 d. *The grandmother of the boys taught themselves to knit.

12 a. The riders injured themselves in the race yesterday.
12 b. *The rider injured themselves in the race yesterday.
12 c. The riders behind the prince injured themselves in the race yesterday.
12 d. *The rider behind the princes injured themselves in the race yesterday.

13 a. The widows supported themselves during the difficult times.
13 b. *The widow supported themselves during the difficult times.
13 c. The widows near the banker supported themselves during the difficult times.
13 d. *The widow near the bankers supported themselves during the difficult times.

14 a. The daughters cleaned themselves up after gardening.
14 b. *The daughter cleaned themselves up after gardening.
14 c. The daughters of the parent cleaned themselves up after gardening.
14 d. *The daughter of the parents cleaned themselves up after gardening.

15 a. The brides prepared themselves for the big party.
15 b. *The bride prepared themselves for the big party.
15 c. The brides behind the hairstylist prepared themselves for the big party.
15 d. *The bride behind the hairstylists prepared themselves for the big party.

16 a. The kids amused themselves during the holidays.
16 b. *The kid amused themselves during the holidays.
16 c. The kids by the lady amused themselves during the holidays.
16 d. *The kid by the ladies amused themselves during the holidays.

17 a. The waiters defended themselves yesterday morning.

17 b. *The waiter defended themselves yesterday morning.
17 c. The waiters by the customer defended themselves yesterday morning.
17 d. *The waiter by the customers defended themselves yesterday morning.

18 a. The teenagers cut themselves by mistake last night.
18 b. *The teenager cut themselves by mistake last night.
18 c. The teenagers by the teacher cut themselves by mistake last night.
18 d. *The teenager by the teachers cut themselves by mistake last night.

19 a. The workers covered themselves with blankets before going to sleep.
19 b. *The worker covered themselves with blankets before going to sleep.
19 c. The workers behind the engineer covered themselves with blankets before going to sleep.
19 d. *The worker behind the engineers covered themselves with blankets before going to sleep.

20 a. The princes surrounded themselves with guards last night.
20 b. *The prince surrounded themselves with guards last night.
20 c. The princes behind the consultant surrounded themselves with guards last night.
20 d. *The prince behind the consultants surrounded themselves with guards last night.

21 a. The boys encouraged themselves to start a volunteering initiative.
21 b. *The boy encouraged themselves to start a volunteering initiative.
21 c. The boys with the manager encouraged themselves to start a volunteering initiative.
21 d. *The boy with the managers encouraged themselves to start a volunteering initiative.

22 a. The patients tested themselves for coronavirus at home.
22 b. *The patient tested themselves for coronavirus at home.
22 c. The patients near the nurse tested themselves for coronavirus at home.
22 d. *The patient near the nurses tested themselves for coronavirus at home.

23 a. The officers prepared themselves for a long night.
23 b. *The officer prepared themselves for a long night.
23 c. The officers with the criminal prepared themselves for a long night.
23 d. *The officer with the criminals prepared themselves for a long night.

24 a. The scientists established themselves independently.
24 b. *The scientist established themselves independently.
24 c. The scientists near the student established themselves independently.
24 d. *The scientist near the students established themselves independently.

Fillers

1. The sister of the bride left the party early last night.
2. The brother of the boy took the black notebook yesterday.
3. The mother by the kids visited the doctor yesterday morning.

4. The tenants behind the landlord lost their keys yesterday.
5. The neighbours by the concierge complained about the noise last night.
6. The chefs with the customers used high-quality ingredients.
7. The clown near the magician enjoyed playing with the kids.
8. The nurse with the doctors mistakenly hurt the patient.
9. The kid with the mother cut his finger with a knife.
10. The managers near the secretary encouraged the employees to learn new skills.
11. The parents behind the teachers educated their children without the government's help.
12. The players by the coach arranged their positions in the game.
13. The daughter did the laundry last week.
14. The mother enjoyed cooking yesterday afternoon.
15. The researcher supported cancer research all of his life.
16. The girl hurt her back when she fell off her horse.
17. The parents taught their children to help others if they can.
18. The manager introduced the new employee to the staff.
19. The man protected his family from any danger.
20. The manager presented the report to his colleagues at the meeting yesterday.
21. The boy forgot his wallet at school yesterday.
22. The businessman visited New York last Summer.
23. The man with the officer defended himself in court.
24. The businessman considered himself lucky.
25. The princess surrounded herself with honest people.
26. The businesswoman organized herself and planned her work in an efficient way.
27. The parents by the doctor are keen to help their children.
28. The engineers never work on the weekends.
29. The coach runs four miles every morning.
30. The teacher has some schoolwork to do.
31. *The pilot by the assistants cannot flying anymore.
32. *The photographer with the lady taking good pictures.
33. *The girls behind the boys is preparing for a surprise party.
34. *The waiters around the chef were had worked at the restaurant for 10 years.
35. *The parents of the kid will visiting the school next week.
36. *The surgeons near the patient are talk about his medical history.
37. *The boys with the teacher will participating in the school's activities.
38. *The doctors near the patient discussing him situation.
39. *The man behind the doctor hurt he while he was walking.
40. *The teacher of the schoolgirl supported she to succeed.
41. *The mother of the kids prepared they for school.
42. * The farmer and the carpenter are enjoy their work in the field.
43. * The housekeepers around the kids are kindest than their teachers.
44. *The student behind the teacher is talls than his brother.
45. *The mother wrote that the child laughed at that slow herself.
46. *The girl said that the teacher laughed at that funny herself.
47. *The taxi driver said that the man hit that careless himself.

48. *The man said that the boy was angry with the lazy himself.
49. *The soldiers know that the generals like today's themselves.
50. *He is one of those mens which always on time.
51. *The women give hers papers to the secretary yesterday.
52. *Her and him friend were in the library yesterday morning.
53. *The student believed that him would succeed.
54. *The nurse thought that herself would leave early today.
55. *John thought that Mary disliked they.
56. *The singers entertained the audience with him performance.
57. *They book was the best seller last month.
58. *The engineer was drives to work.
59. * The boy was walks to the store and bought milk every morning.
60. *The girl by the boy is looked beautiful.

Appendix E

A)

1 a. The students quickly organized themselves after entering the classroom yesterday.
1 b. *The student quickly organized themselves after entering the classroom yesterday.
Q: Were/was the student(s) in the classroom yesterday? Yes.

2 a. The travellers completely isolated themselves for ten days after arrival.
2 b. *The traveller completely isolated themselves for ten days after arrival.
Q: Was the isolation for fifteen days? No.

3 a. The boys apparently protected themselves from illness by healthy eating.
3 b. *The boy apparently protected themselves from illness by healthy eating.
Q: Did the boy(s) eat healthy food? Yes.

4 a. The girls obviously covered themselves with warm clothing earlier today.
4 b. *The girl obviously covered themselves with warm clothing earlier today.
Q: Did the girl(s) wear light clothing? No.

5 a. The actors quickly prepared themselves for the cameras last night.
5 b. *The actor quickly prepared themselves for the cameras last night.
Q: Were/was the actor(s) prepared for the cameras last night? Yes.

6 a. The students really pushed themselves to finish the long assignment.
6 b. *The student really pushed themselves to finish the long assignment.
Q: Was the assignment short? No.

7 a. The schoolgirls clearly trained themselves to dance on a stage.
7 b. *The schoolgirl clearly trained themselves to dance on a stage.
Q: Were/was the schoolgirl(s) good at dancing? Yes.

8 a. The girls quickly dressed themselves before going to school yesterday.
8 b. *The girl quickly dressed themselves before going to school yesterday.
Q: Were/was the girl(s) slow in getting dressed? No.

9 a. The students happily gave themselves a break after the class.
9 b. *The student happily gave themselves a break after the class.
Q: Was there a class before the break? Yes.

10 a. The boys carefully washed themselves using soap before having dinner.
10 b. * The boy carefully washed themselves using soap before having dinner.
Q: Did the boy(s) have dinner before washing? No.

11 a. The managers apparently forced themselves to complete the unfinished projects.
11 b.*The manager apparently forced themselves to complete the unfinished projects.
Q: Were there any projects that need to be finished? Yes.

12 a. The students surprisingly hurt themselves after failing the final exam.
12 b. *The student surprisingly hurt themselves after failing the final exam.
Q: Did the student(s) do well in the final exam? No.

B)

1 a. The girls by the boys accidently hurt themselves while playing in the garden.
1 b. The girls by the boy accidently hurt themselves while playing in the garden.
1 c. *The girl by the boys accidently hurt themselves while playing in the garden.
1 d. *The girl by the boy accidently hurt themselves while playing in the garden.
Q: Did somebody get hurt while playing in the garden? Yes

2 a. The teachers of the students likely found themselves frustrated with the low grades.
2 b. The teachers of the student likely found themselves frustrated with the low grades.
2 c. *The teacher of the students likely found themselves frustrated with the low grades.
2 d. *The teacher of the student likely found themselves frustrated with the low grades.
Q: Were the grades high? No

3 a. The boys behind the officers definitely caused themselves a great deal of trouble.
3 b. The boys behind the officer definitely caused themselves a great deal of trouble.
3 c. *The boy behind the officers definitely caused themselves a great deal of trouble.
3 d. *The boy behind the officer definitely caused themselves a great deal of trouble.
Q: Was somebody in trouble? Yes

4 a. The assistants opposite the researchers carefully prepared themselves for the big research project.
4 b. The assistants opposite the researcher carefully prepared themselves for the big research project.
4 c. *The assistant opposite the researchers carefully prepared themselves for the big research project.
4 d. *The assistant opposite the researcher carefully prepared themselves for the big research project.
Q: Was the project small? No

5 a. The teenagers near the parents happily entertained themselves during the long car journey.
5 b. The teenagers near the parent happily entertained themselves during the long car journey.
5 c. *The teenager near the parents happily entertained themselves during the long car journey.
5 d .*The teenager near the parent happily entertained themselves during the long car journey.

Q: Was the journey long? Yes

6 a. The girls by the boys unsurprisingly blamed themselves for the horrible car accident.
6 b. The girls by the boy unsurprisingly blamed themselves for the horrible car accident.
6 c. *The girl by the boys unsurprisingly blamed themselves for the horrible car accident.
6 d. *The girl by the boy unsurprisingly blamed themselves for the horrible car accident.
Q: Was it a minor car accident? No

7 a. The students of the teachers positively encouraged themselves to learn something new daily.
7 b. The students of the teacher positively encouraged themselves to learn something new daily.
7 c. * The student of the teachers positively encouraged themselves to learn something new daily.
7 d. * The student of the teacher positively encouraged themselves to learn something new daily.
Q: Was somebody excited to learn new things? Yes

8 a. The mothers behind the girls clearly enjoyed themselves walking in the sunny weather.
8 b. The mothers behind the girl clearly enjoyed themselves walking in the sunny weather.
8 c. *The mother behind the girls clearly enjoyed themselves walking in the sunny weather.
8 d. *The mother behind the girl clearly enjoyed themselves walking in the sunny weather.
Q: Was it raining? No

9 a. The protesters opposite the officers quickly surrounded themselves with a group of supporters.
9 b. The protesters opposite the officer quickly surrounded themselves with a group of supporters.
9 c. *The protester opposite the officers quickly surrounded themselves with a group of supporters.
9 d. *The protester opposite the officer quickly surrounded themselves with a group of supporters.
Q: Was there a group of supporters? Yes

10 a. The guests near the hosts definitely served themselves food from the buffet table.
10 b. The guests near the host definitely served themselves food from the buffet table.
10 c. * The guest near the hosts definitely served themselves food from the buffet table.
10 d. * The guest near the host definitely served themselves food from the buffet table.
Q: Was there anybody serving food to others? No

11 a. The secretaries by the directors probably introduced themselves to the new candidates yesterday.
11 b. The secretaries by the director probably introduced themselves to the new candidates yesterday.
11 c. * The secretary by the directors probably introduced themselves to the new candidates yesterday.
11 d. * The secretary by the director probably introduced themselves to the new candidates yesterday.
Q: Were there any new candidates? Yes

12 a. The photographers of the actors successfully marketed themselves using popular social media sites.
12 b. The photographers of the actor successfully marketed themselves using popular social media sites.
12 c. *The photographer of the actors successfully marketed themselves using popular social media sites.
12 d. *The photographer of the actor successfully marketed themselves using popular social media sites.
Q: Do/Does the actor(s) use social media for marketing? No

13 a. The guides opposite the tourists apparently taught themselves about different cultures and backgrounds.
13 b. The guides opposite the tourist apparently taught themselves about different cultures and backgrounds.
13 c. *The guide opposite the tourists apparently taught themselves about different cultures and backgrounds.
13 d. *The guide opposite the tourist apparently taught themselves about different cultures and backgrounds.
Q: Was/Were the guide(s) opposite the tourist(s)? Yes

14 a. The housekeepers behind the ladies clearly amused themselves by playing cards and games.
14 b. The housekeepers behind the lady clearly amused themselves by playing cards and games.
14 c. *The housekeeper behind the ladies clearly amused themselves by playing cards and games.
14 d. *The housekeeper behind the lady clearly amused themselves by playing cards and games.
Q: Was the farmer playing cards? No

15 a. The waiters near the chefs apparently prepared themselves for the big dinner party.
15 b. The waiters near the chef apparently prepared themselves for the big dinner party.
15 c. *The waiter near the chefs apparently prepared themselves for the big dinner party.
15 d.*The waiter near the chef apparently prepared themselves for the big dinner party.
Q: Was there a dinner party? Yes

16 a. The nurses of the patients safely protected themselves from infections by wearing masks.
16 b. The nurses of the patient safely protected themselves from infections by wearing masks.
16 c. * The nurse of the patients safely protected themselves from infections by wearing masks.
16 d. * The nurse of the patient safely protected themselves from infections by wearing masks.
Q: Was/Were the patient(s) wearing masks? No

17 a. The soldiers by the journalists unsurprisingly placed themselves in the line of fire.
17 b. The soldiers by the journalist unsurprisingly placed themselves in the line of fire.
17 c. * The soldier by the journalists unsurprisingly placed themselves in the line of fire.
17 d. * The soldier by the journalist unsurprisingly placed themselves in the line of fire.
Q: Was/Were the journalist(s) close to the soldier(s)? Yes

18 a. The nurses behind the doctors carefully prepared themselves to fight against infectious diseases.
18 b. The nurses behind the doctor carefully prepared themselves to fight against infectious diseases.
18 c. *The nurse behind the doctors carefully prepared themselves to fight against infectious diseases.
18 d. *The nurse behind the doctor carefully prepared themselves to fight against infectious diseases.
Q: Was/Were the doctor(s) behind the nurse(s)? No

19 a. The participants opposite the researchers happily helped themselves to some sweets earlier today.
19 b. The participants opposite the researcher happily helped themselves to some sweets earlier today.
19 c. *The participant opposite the researchers happily helped themselves to some sweets earlier today.
19 d. * The participant opposite the researcher happily helped themselves to some sweets earlier today.
Q: Were there some sweets? Yes

20 a. The engineers near the reporters confidently introduced themselves by talking about some achievements.
20 b. The engineers near the reporter confidently introduced themselves by talking about some achievements.
20 c. *The engineer near the reporters confidently introduced themselves by talking about some achievements.
20 d. *The engineer near the reporter confidently introduced themselves by talking about some achievements.
Q: Did the reporter(s) talk about some achievements? No

21 a. The managers by the assistants always supported themselves despite what had happened earlier.
21 b. The managers by the assistant always supported themselves despite what had happened earlier.
21 c. *The manager by the assistants always supported themselves despite what had happened earlier.
21 d. *The manager by the assistant always supported themselves despite what had happened earlier.
Q: Had something happened earlier? Yes

22 a. The singers behind the musicians probably taught themselves to play different musical instruments.
22 b. The singers behind the musician probably taught themselves to play different musical instruments.

22 c. *The singer behind the musicians probably taught themselves to play different musical instruments.
22 d. *The singer behind the musician probably taught themselves to play different musical instruments.
Q: Was/Were the musician(s) behind the singer(s)? No

23 a. The athletes opposite the coaches obviously injured themselves while training and playing football.
23 b. The athletes opposite the coach obviously injured themselves while training and playing football.
23 c. *The athlete opposite the coaches obviously injured themselves while training and playing football.
23 d. *The athlete opposite the coach obviously injured themselves while training and playing football.
Q: Was/were the athlete(s) playing football? Yes

24 a. The boys near the girls accidently burnt themselves while cooking in the kitchen.
24 b. The boys near the girl accidently burnt themselves while cooking in the kitchen.
24 c. *The boy near the girls accidently burnt themselves while cooking in the kitchen.
24 d. *The boy near the girl accidently burnt themselves while cooking in the kitchen.
Q: Did the girl(s) get hurt? No

Fillers

1.The friend of the girl surprisingly was late for the exam yesterday.

Q: Was somebody late for the exam? Yes.

2. The pilot near the assistants clearly was tired after the long flight.

Q: Was it a short flight? No.

3. The actor behind the ladies apparently is happy about his performance.

Q: Was the actor happy? Yes.

4. The waiter with the chef is friendly but professional.

Q: Was the waiter unfriendly? No.

5. The children behind the woman surprisingly were afraid of clowns yesterday.

Q: Were the clowns scary to the children? Yes.

6. The scientists near the librarian obviously are happy about the results.

Q: Was somebody unhappy? No.

7. The parents of the kids were completely involved in everything their kids did.

Q: Were the parents engaged in their kids' lives? Yes.

8. The boys behind the man are trained to be football players.

Q: Are the boys expected to be basketball players? No.

9. The businessman by the secretary probably donated to the charity he supports.

Q: Does the businessman support a particular charity? Yes.

10. The girl with the gardener happily planted many flowers in the backyard.

Q: Did the girl plant vegetables? No.

11. The workers near the engineer certainly worked on dangerous construction sites.

Q: Was there any engineer with the workers? Yes.

12. The girls by the chef quickly cut the potatoes into squares before boiling.

Q: Were the potatoes cut into slices? No.

13. The farmers behind the housekeepers apparently worked hard for many hours in the field.

Q: Did the farmers work for long hours? Yes.

14. The boys behind the policeman probably injured our old neighbour.

Q: Was the policeman behind the boys? No.

15. The students near the teachers taught their classmates how to solve some maths problems.

Q: Did the students help their classmates with maths? Yes.

16. The kids with the farmers apparently hurt the animals there.

Q: Were the kids with their parents? No.

17. The dentist near the nurse marketed his dental clinic on social media.

Q: Was the clinic advertised on social media? Yes.

18. The consultant by the director surprisingly convinced the clients to buy expensive products.

Q: Did the clients buy cheap products? No.

19. The painter with the woman confidently described himself as an international artist.

Q: Was the painter confident? Yes.

20. The sister of the boys probably hid herself under the blankets.

Q: Was the girl hiding under the table? No.

21. The woman with the doctor surprisingly introduced herself in Spanish.

Q: Does the woman speak Spanish? Yes.

22. The dog of the girl was left in the house by itself for a week.

Q: Does the dog belong to a boy? No.

23. The horse with the boy sadly hurt itself yesterday in the race.

Q: Was the horse in pain? Yes.

24. The boy near the nurse mistakenly pricked himself with a used syringe needle.

Q: Was the needle clean? No.

25. After presenting themselves and their work, they asked for some suggestions and feedback.

Q: Did they ask for feedback after the presentation? Yes.

26. Before introducing themselves, they asked the audience some questions about the project.

Q: Did the audience ask questions about the project? No

27. After preparing themselves, they decided to visit more than one place.

Q: Were they prepared to visit many places? Yes.

28. Before isolating themselves, they shopped for fresh food and canned vegetables.

Q: Did they shop for food after isolation? No.

29. Our neighbours had the same suitcase as us and we accidently picked theirs.

Q: Were the suitcases identical? Yes.

30. The neighbours have built a huge fence between our house and theirs.

Q: Was the fence short? No.

31. My sisters and I had different ideas for Mother's Day but their ideas are better.

Q: Did the sisters have ideas for Mother's Day? Yes.

32. The neighbours got a new brand car and my father plans to buy one like theirs.

Q: Did the neighbours buy a used car? No.

33. The boy did all his homework, but the girl did not do hers.

Q: Did the girl have homework? Yes.

34. The boy did not have an umbrella, so the mother lent him hers.

Q: Did the boy lend his mother an umbrella? No.

35. The husband and wife bought black cars but the bigger one is his.

Q: Does the husband have the bigger car? Yes.

36. The man who bought a brand new car obviously is old.

Q: Did the man buy a used car? No.

37. The nurse met all her patients, but the doctor did not meet his.

Q: Were there any patients? Yes.

38. The secretary who sent all the documents clearly is professional.

Q: Was the secretary unprofessional? No.

39. The author who wrote the book certainly was happy with the good reviews.

Q: Did the reviews make the author happy? Yes.

40. The students that the teacher met apparently are very competitive.

Q: Did the students meet their friends? No.

41. The lady who gave the speech apparently was the department head.

Q: Was the department head a lady? Yes.

42. The boys that the police fined evidently breached the rules.

Q: Did the boys follow the rules? No.

43. The patients that the doctor examined were severely ill.

Q: Were the patients very ill? Yes.

44. The engineers that the waiter greeted recently started a new business.

Q: Did the waiter start a new business? No.

45. The little girl whose doll was lost clearly was sad.

Q: Was somebody unhappy? Yes.

46. The man whose daughter is a doctor lives in London.

Q: Is the man's son a doctor? No.

47. The woman whose name is Greek is the guest speaker.

Q: Was the guest speaker a woman? Yes.

48. The author whose book is popular is a doctor.

Q: Is the author an engineer? No.

49. The manager is the man who your father met at our house last week.

Q: Is the manager a man? Yes.

50. The teacher is the lady who Sara invited to the party last weekend.

Q: Was the teacher invited to a conference? No.

51. This is the patient that the nurse spent all night with.

Q: Was there a nurse with the patient? Yes.

52. This is the new employee that the manager introduced yesterday.

Q: Was the employee introduced by the secretary? No.

53. The man was walking back to his flat when he saw the accident.

Q: Does the man live in a flat? Yes.

54. The mother was cooking the lunch when the child fell down the stairs.

Q: Was the mother watching TV? No.

55. The girl was watching TV when the neighbours visited.

Q: Were there any visitors? Yes.

56. The boy was playing in the backyard when the mother left the house.

Q: Was the mother in the backyard with the boy? No.

57. The woman heard the baby cry as she was cleaning the house.

Q: Does the woman have a baby? Yes.

58. The man heard the phone ringing as he was reading the newspapers.

Q: Was the man sleeping? No.

59. The boy lost his umbrella as he was walking to school.

Q: Did the boy lose something? Yes.

60. The girl hurt her leg as she was playing in the park.

Q: Did the girl hurt her arm? No.

61. The man was interviewed for a teaching position.

Q: Was the interview for a teaching job? Yes.

62. Data was gathered through questionnaires.

Q: Was the data gathered through interviews? No.

63. The patient was visited this morning.

Q: Did the patient have any visitors this morning? Yes.

64. The house was decorated for Christmas.

Q: Was the house decorated for a wedding party? No.

65. Only the girls are allowed to attend the party.

Q: Can the girls attend the party? Yes.

66. The boy apparently went to the museum yesterday.

Q: Did the boys go to the park yesterday? No.

67. The girl happily baked some cookies this afternoon.

Q: Did the girl bake something this afternoon? Yes.

68. The dog surprisingly covered itself with dirt.

Q: Was the dog clean? No.

69. The man unfortunately missed the train to London.

Q: Did the man want to go to London? Yes.

70. The girl always puts herself in situations where she could be at risk.

Q: Does the girl always know how to stay safe? No.

71. The minister proudly introduced himself at the Italian restaurant in town.

Q: Was the minister in an Italian restaurant? Yes.

72. The students were supported in using the library.

Q: Were the students ignored in the library? No.

73. The schoolboy has clearly forgotten his notebook.

Q: Did the schoolboy forget something? Yes.

74. The sisters have lived in France since 2000.

Q: Were the sisters living in France in 1990? No.

75. The boy unfortunately missed the classes yesterday.

Q: Was the boy absent yesterday? Yes.

76. Sara has successfully finished her internship at the museum.

Q: Was Sara's internship at the hospital? No.

77. The mother accidently cut herself with a kitchen knife yesterday.

Q: Did the mother use a knife yesterday? Yes.

78. The boy quickly prepared himself to meet his friends.

Q: Was the boy prepared to meet his teachers? No.

79. The lady surprisingly hurt herself last night in the kitchen.

Q: Was somebody hurt last night? Yes.

80. The baby surprisingly helped himself to drink his bottle.

Q: Was the baby eating food? No.

81. Jane and Sara gave each other gifts.

Q: Did Sara get a gift from Jane? Yes.

82. Jack and Dan were talking to each other in the cafeteria.

Q: Was Jack talking to his mother in the cafeteria? No.

83. The sisters call each other regularly.

Q: Are the sisters staying in touch? Yes.

84. The chefs at the restaurant met each other ten years ago.

Q: Is there only one chef at the restaurant? No.

85. The managers refused the plans suggested by the coordinator.

Q: Did the coordinator suggest any plans? Yes.

86. The teachers liked the slides presented by the student.

Q: Did the teachers present the slides? No.

87. The directors approved the project introduced by the engineer.

Q: Did the engineer introduce the project? Yes.

88. The father paid for the course needed by his daughter.

Q: Did the university pay for the course? No.

89. The parents appreciated the efforts made by their son.

Q: Were the parents happy about their son? Yes

90. The mother liked the dinner prepared by the girl.

Q: Did the mother prepare the dinner? No.