MORPHOLOGICAL AND IDENTITY PRIMING IN WORD LEARNING AND TEXT READING AS A WINDOW INTO THE MENTAL LEXICON

MORPHOLOGICAL AND IDENTITY PRIMING IN WORD LEARNING AND TEXT READING AS A WINDOW INTO THE MENTAL LEXICON

By MELDA COSKUN, MSc

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2022) Hamilton, Ontario

(Cognitive Science of Language)

TITLE: Morphological and Identity Priming in Word Learning and Text Reading as a Window into the Mental Lexicon

AUTHOR: Melda Coskun, M.Sc. (Middle East Technical University), B.Sc. (Baskent University) SUPERVISOR: Dr. Victor Kuperman

NUMBER OF PAGES: xi, 141

To my family,
for their unconditional love..

Abstract

This thesis examines the influence of morphological and identity priming to understand how repetition influences word recognition and novel word learning in first (L1) and second (L2) language adults. The following questions are addressed: How does morphological relatedness between repeated words influence (i) word recognition in natural reading and (ii) novel word learning? (iii) What interactions exist between word repetition and selective attention in novel word learning?

Chapter 2 addresses question (i), finding little evidence of morphological priming effects (i.e., faster recognition of a word following a morphologically related word) in L2 reading, and none in L1. The effects of identity priming were ubiquitous in both groups.

Chapter 3 examines question (ii) for L1 readers. Low-frequency base words (e.g., *caltrop*) and novel complex forms (e.g., *caltroper*) of those bases were primed by two repetitions of identical forms or alternate forms. Learning performance was consistently as good or better after identity priming than after morphological priming. However, orthographic and semantic learning for base forms was stronger in the morphological priming condition.

Chapter 4 examines question (iii). Attention was manipulated by delivering attention-inducing instructions, while the control group received no instructions. Exposure was manipulated by embedding novel words either 2, 4, or 8 times. The presence of instruction led to a short-lived speed-up in eye-movements and faster recognition of novel words. Critically, L1 learners reached optimal performance in the post-tests earlier (after 4 exposures), while L2 learners' performance continued to improve through more exposures.

Overall, this thesis shows that morphological priming facilitated L2 visual word recognition and L1 novel word learning when a complex form is a prime, and the base form is a target. We discuss reasons for this asymmetric effect and these results in the framework of the theories of word learning and morphological processing.

Acknowledgments

As I have come to the end of this seemingly endless journey, I would like to thank everyone who has helped me along the way and played a significant role in my success.

First and foremost, I would like to thank my dear supervisor, Dr. Victor Kuperman. Victor, I consider myself extremely fortunate to have accomplished this journey under your guidance. Your immense knowledge and extensive experience have played a huge role in this success. Thank you for all of the opportunities you have provided me, through which I have been able to improve both my technical and soft abilities. Besides being a great scientist, you are also a person with a big heart. When I needed it the most, your kindness and outstanding empathy abilities have always warmed my heart.

I would also like to thank my committee members. Dr. Anna Moro and Dr. Daniel Pape. Your feedback in our meetings has been very helpful in the development of these projects. Thank you for all of your contributions.

I would also like to express my gratitude to Dr. Lili Service and Dr. John Connolly for their guidance during my comprehensive project. I would like to extend my sincere thanks to Dr. Jay Rueckl, for his assistance and advice during the development of the project we are working on together.

I would also like to thank my dear professors who supported me in pursuing my doctorate abroad. Dr. Cengiz Acarturk, Dr. Umut Ozge, Dr. Duygu Ozge and Dr. Annette Hohenberger, thank you for your support during the application process.

Thank you to all of the Reading Lab members who made this journey enjoyable and productive in every way. I would like to express my gratitude to both former and current Reading Lab members whose paths cross mine. I would like to thank Kaitlyn Battershill, Megan Karabin, and Heather Wild in particular for their invaluable ideas and suggestions during the completion of this thesis. A special thanks to Nadia Lana, with whom we collaborated on one of the projects. Thank you for your intellectual and emotional support. I would also want to express my heartfelt gratitude to Gaisha Oralova for being a wonderful friend and colleague to me throughout the years.

Finally, I want to thank my family. First of all, I would like to thank my dear parents, Ayse and Haluk. Thank you for compromises you made for my education. Thank you for always encouraging me to think critically and creatively. Thank you for making me an individual who believes that I can achieve what I believe in. I consider myself extremely fortunate to be your daughter. I would also like to thank my siblings, Dilara and Baris. You are my light, my motivation, and the reason I keep convincing myself to do better. Thank you! .. and my dear husband Firat, you took on all the stress of this journey as much as I did. Thank you for sharing my dreams and encouraging me to pursue them. Thank you for always being there for me when I needed you, with care and affection. Finally, I would like to commemorate my dear uncle Bekir, who passed

away in 2020.. You were outstanding in your ability to use words in such a heartfelt manner. I wish you could have been here to see I did a PhD on *words*. I dedicate this thesis to my family, whose unconditional love has nurtured me and made me become the person I am today.

Table of Contents

Introduction 1

Long-lag repetition priming in natural text reading: Little to no evidence for morphological effects. 20

- Abstract 20
- Introduction 21
- Results 31
- General Discussion...37
- References 42
- Supplementary Materials..50

The Role of Morphological Priming in Orthographic and Semantic Word Learning 56

- Abstract 56
- Introduction 57
- Methods..62
- Results 66
- General Discussion...70
- References 72
- Supplementary Materials..77

Exposure and Attention in Novel Word Learning in the First and Second Language... 84

- Abstract 84
- Introduction 85
- Results and Discussion..94
- General Discussion...103
- References 107
- Supplementary Materials..112

Conclusion 124

List of Tables

Chapter 2

Table 1. Summary of eye-movement characteristics in the three subcorpora.. 33

Table 2. Estimated identity (*bf1-bf2* and *if1-if2*) and inflection (*bf1-if2* and *if1-bf2*) priming effects.. 35

Table S1. Results of the linear mixed-effects models fitted to Total Fixation Times for English L1, Dutch L1, and English L2.. 50

Table S2. Results of the linear mixed-effects models fitted to Gaze Durations for English L1, Dutch L1, and English L2 51

Table S3. Results of the generalized linear mixed-effects models fitted to Regressions for English L1, Dutch L1, and English L2 52

Table S4. Results of the generalized linear mixed-effects models fitted to Skips for English L1, Dutch L1, and English L2... 53

Table S5. Conditional means estimated based on regression models for the six levels of factor Type, i.e., the six categories formed by the type of the repeated word (base form *bf* of inflected form *if*), and its position in the pair of occurrences (first or second) 54

Chapter 3

Table 1. The experimental conditions tested 63

Table 2. Example paragraphs. Sentences differing between conditions are italicised 64

Table 3. The distribution of the frequencies of the given answers based on the Target Type of the recalled word form (identity and morphological priming conditions jointly)... 70

Table S1. The list of base words with their definitions. 77

Table S2.The list of prime and target words used in the experiments, and the paragraphs in which these words were embedded 77

Table S3. Summary of Orthographic Choice, Definition Matching, and Definition Prompting Tests Results... 82

Table S4.Summary of the generalized linear mixed effect regression models fitted to the accuracy and RTs of orthographic choice task, the accuracy of definition matching task, and the definition prompting exact. 83

Chapter 4

Table S1. Participant self-ratings in the English proficiency skills. 112

Table S2.Total novel word reading times (ms) in the L1 and L2 113

Table S3.ANOVA summary of a model fitted to total fixation time in L1 participants 113

Table S4. ANOVA summary of a model fitted to total fixation time in L1 participants 113

Table S5. Summary of the generalized linear mixed effect regression model fitted to the accuracy of orthographic choice task in L1 participants............... 114
Table S6. Summary of the generalized linear mixed effect regression model fitted to the accuracy of orthographic choice task in L2 participants............... 114
Table S7. Summary of the generalized linear mixed effect regression model fitted to RTs of orthographic choice task in L1 participants............................ 114
Table S8. Estimated marginal means contrasts of Instruction per level of Session for the L1 readers.. 115
Table S9. Summary of the generalized linear mixed effect regression model fitted to RTs of orthographic choice task in L2 participants............................ 115
Table S10. Summary of the generalized linear mixed effect regression model fitted to definition matching scores in L1 participants 115
Table S11. Summary of the generalized linear mixed effect regression model fitted to definition matching scores in L2 participants 116
Table S12. Summary of the generalized linear mixed effect regression model fitted to definition prompting scores in L1 participants.......................... 116
Table S13. Summary of the generalized linear mixed effect regression model fitted to definition prompting scores in L2 participants.......................... 116

List of Figures

Chapter 3

Figure 1. Partial effects of Priming Type and Target Type on orthographic decision accuracy (left) and RT (right) 68

Figure 2. Partial effects of Priming Type and Target Type on definition matching scores (left) and definition prompting scores (right). 68

Chapter 4

Figure 1. L1 (left) and L2 (right) total fixation time by conditions of Instruction and Exposure 96

Figure 2.Total fixation times to every occurrence of the target in L1 (left) and L2 (right) data, by conditions of Exposure and Instruction............................ 97

Figure 3. L1(left) and L2 (right) Orthographic choice accuracy scores per Session broken down per level of Exposure 99

Figure 4. L1(left) and L2 (right) Estimated values of RT per Session and Instruction 100

Figure 5. L1 (left) and L2 (right) Definition matching scores per Instruction broken down per level of Exposure 101

Figure 6. L1 (left) and L2 (right) Definition prompting scores per level of Exposure... 103

Declaration of academic achievement

This is a "sandwich" thesis, as defined by McMaster University's School of Graduate Studies. It contains three empirical investigations on which I am the primary author. The following sections detail the roles of each author for the three studies in this dissertation.

Chapter 2

This study has been submitted and currently under review in *The Mental Lexicon* as **Coskun**. M., Kuperman. V., Rueckl. J. (submitted). Long-lag repetition priming in natural text reading: Little to no evidence for morphological effects.

Coskun, M: study design, literature review, eye-tracking data analysis, manuscript writing, revision, and preparation for publication.

Kuperman, V: study design, eye-tracking data analysis, manuscript writing and revision.

Rueckl, J: manuscript revision.

Chapter 3

This study has been submitted and currently under review in *Language Learning and Development* as **Coskun**, M., Kuperman, V. (submitted). The Role of Morphological Priming in Orthographic and Semantic Word Learning.

Coskun, **M**: study design, stimuli preparation, literature review, data collection, eye-tracking data analysis, manuscript writing, revision, and preparation for publication.

Kuperman, V: study design, eye-tracking data analysis, manuscript writing and revision.

Chapter 4

This chapter has been submitted and currently under review in *Journal of Cognitive Psychology* as Coskun, M., Lana, N., Kuperman, V., Exposure and Attention in Novel Word Learning in the First and Second Language

Coskun, M: study design, stimuli preparation, experiment programming, literature review, data collection, eye-tracking data analysis, manuscript writing, revision, and preparation for publication.

Lana, N., study design, stimuli preparation, literature review, data collection, manuscript writing, revision, and preparation for publication.

Kuperman, V: study design, eye-tracking data analysis, manuscript writing and revision.

Additional achievements

In addition to the studies presented in this dissertation, the author contributed to the publication of the following paper:

Acarturk, C., **Coskun, M**., & Emil, S. (2022). Multimodal communication in instructional settings: An investigation of the functional roles of gestures and arrows. *Revista Signos. Estudios de Lingüística*, *54*(107).

Introduction

A typical English speaker is thought to know more than 70,000 words (Brysbaert, Stevens, Mandera, & Keuleers, 2016). This means that, starting in their first year of life, one can learn around 10 new words per day to achieve that level of vocabulary. When other languages that one may be familiar with are considered, this number grows even larger. The human mind not only has a strong capacity for comprehending and using the words already known, but it also has a remarkable ability to understand words never seen before, to create new words instantly by modifying known ones, and, of course, to learn new words continuously. The question then becomes, how does the human mind learn such a large number of words and access them in a seamless and effortless manner?

This question is at the core of mental lexicon research and raises many additional questions such as: What are the characteristics of lexical information stored in the lexicon; how are words accessed during language production or comprehension; how are new words added to mental lexicon; what are the underlying mechanisms that enable lexical activities such as word recognition and word learning; how do languages, tasks, and individuals affect the organization and functioning of the mental lexicon (see Libben & Jarema, 2002, for a review).

This thesis provides an empirical investigation of a few issues in mental lexicon research that have previously been explored in a specific or limited manner, by introducing new dimensions (such as adding new conditions or factors of interest) and providing novel methodological approaches to the issues. By varying the word-internal (e.g., morphological relatedness vs identity) or word-external (exposure frequency, attention to the exposures) variables of the repeated words, this thesis investigated the influence of repetition on word learning and word recognition in native (L1) and non-native (L2) adult readers. These issues are summarized by the research questions below.

1. How does morphological priming (i.e., faster recognition of a word preceded by a morphologically related word) influence word recognition in natural reading of long texts in L1 and L2 readers?
2. How does morphological priming affect novel word learning in adult native readers?
3. How do the well-known predictors of word learning – i.e., exposure and selective attention – interact in their effects on novel word learning for both L1 and L2 readers? How many exposures to a novel word leads to a sufficient or to a maximum semantic or orthographic knowledge for both groups?

We believe that this thesis will contribute to mental lexicon research on issues such as how words are organized and accessed in the mental lexicon, the underlying mechanisms that enable novel word learning and word processing, and factors that influence novel word learning.

The following sections will first introduce basic concepts and then explain how these research questions will be addressed in this thesis.

The mental lexicon

In early psycholinguistic and linguistic theories, the mental lexicon was likened to a dictionary in the mind (Aitchison, 2012; see also Dóczi, 2019; Libben, Goral, & Libben, 2017; for a review). From this perspective, the mental lexicon was seen as a passive repository storing the specific properties of a word such as its semantic, phonological, and morphological information. The action that produces a linguistic outcome, on the other hand, would be generated by a set of grammatical rules that are not part of the lexicon but are supposed to operate on these stored representations (Chomsky, 1965, also see Dóczi, 2019; Kuperman, Jarema, & Libben, 2021; Libben, Schwieter, 2019 for a review). However, there are a few issues with this approach, since words cannot be isolated and passive components of the language system (Libben, 2019). Human lexical knowledge, perhaps more than any other component of the human language system, is the one that changes the most and constantly over the course of a lifetime. People consistently add new words to their mental lexicon (Brysbaert et al., 2016; Verhaeghen, 2003); words which are not used frequently or recently can be forgotten; the meanings attributed to words may expand or change; new words can be created instantly by a language user, and understood by the audience despite having never heard it before. These examples demonstrate the dynamic nature of the mental lexicon. Furthermore, lexical representations are not isolated *things*, but rather a part of a highly interconnected system in which a change in one lexical item might have a significant impact on the entire system (Kuperman, Jarema, & Libben, 2021). As a result, the psycholinguistic construct of a mental lexicon has evolved into a dynamic and integrated system. These aspects are highlighted by Jarema and Libben (2007), who define the mental lexicon as "the cognitive system that constitutes the capacity for conscious and unconscious lexical activity". According to this definition, every activity involving words, such as writing, singing, and texting, and also novel word learning and text reading, is considered to be part of the mental lexicon.

Defining a word

Defining what a word is can be difficult, since definitions vary considerably depending on the assumptions that underpin them. For example, for many people defining a word can be as simple as this: When a linguistic unit is isolated from other linguistic units on both sides by spaces, it is called a word. This straightforward orthographic definition is certainly beneficial in covering a number of instances in many languages. However, just as not every language has a writing system (e.g., the indigenous languages of the Americas), space between words is not a universal convention in all writing systems (e.g., Chinese). Even for writing systems that meet these criteria, it is still unclear how some situations (described below) should be treated under this definition.

Consider the following examples to illustrate a few of these scenarios: When a "word" has multiple forms, such as *tree* and *trees*, are they the same word or two different words? Can we say that a word with multiple meanings, such as *round* in *round eyes* and *all year round*, is still the same word? Should we count each individual item separately or the whole phrase as a single word, when multiple linguistic units, such as compound words or idioms, consist of more than one item but express only one meaning (e.g., *bus driver*)? To tackle these issues, Carter (2012) proposes using the term *lexical item*, which acknowledges all possible orthographic, semantic, phonological, and grammatical differences and treats each of these as a separate lexical item.

Nation (1990) approaches the subject from the standpoint of language acquisition. He emphasizes the importance of learners being aware of the criterion that distinguishes one word from another. From this perspective, two lexical items being considered the same word is contingent on whether any additional learning is required in order to learn the second lexical item. For instance, if a learner already knows the past tense rule and one of the verbs *walk* or *walked*, these two items can be treated as the same word because the existing knowledge makes both forms of the verb available.

One other approach, which is commonly used in corpus studies, is to consider words as *tokens* and *types*. In that case, *tokens* are individual linguistic units separated by spaces, and each unique token is referred to as a *type*. For example, the phrase *my cat, little cat, lovely cat* has six *tokens* (*my, cat, little, cat, lovely, cat*) and four *types* (*my, cat, little, lovely*).

Lemma and *word forms* are the terms used in the language acquisition domain. A *lemma* (*base form*) refers to a dictionary form of a word, like the infinitive of a verb (e.g., *walk*), and the nominative singular of a noun (e.g., *house*). *Word forms* refer to the base and its inflected and derived forms.

For the rest of this thesis the term *word* will be used for general reference. The terms *type*, *tokens*, *word forms*, and *base form* will be used if further specifications are needed. The term *lexical item* will be used as a "neutral hold-all term*"* (Carter, 2012).

1. How does long-term morphological priming affect word recognition in the natural reading of long texts?

In many languages, including English, words with more than one morpheme (e.g. *driver*, *blackbird, smaller*) form the majority (Libben, Goral, & Baayen, 2017). Therefore, unsurprisingly, over a half-decade of research has focused on how these complex words are stored, learned, and retrieved during language production or comprehension (Baayen et al., 1997; Baayen, Milin, Đurđević, Hendrix, & Marelli, 2011; Frauenfelder & Schreuder, 1992; Giraudo & Grainger, 2001; Harm & Seidenberg, 1999, 2004; Plaut & Gonnerman, 2000; Schreuder & Baayen, 1995; Seidenberg & Gonnerman, 2000; Taft, 2004). Despite their numerous disagreements on other topics (as will be discussed below), almost all empirical evidence and theoretical accounts agree that when two morphologically related words are presented sequentially, whether visually or auditorily, the effort required to recognize the second word is dramatically reduced as a result of the presence of this morphological relationship. This effect is known as the morphological priming effect (e.g., Münte, Say, Clahsen, Schiltz, & Kutas, 1999; Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997; Raveh & Rueckl, 2000; Stanners, Neiser, Hernon, & Hall, 1979). However, much of the evidence for this conclusion comes from paradigms in which words are presented in isolation. Therefore, the question of whether the data collected from context-less presentation of isolated words can be generalized over natural text reading remains to be answered. One of the objectives of this study is to attempt to answer this question.

First, I will provide a brief background on theoretical models of morphological representations and processing that are primarily relevant to the first and (to a lesser extent) second objectives of the thesis. The majority of the reviews here are based on Diependaele, Grainger, and Sandra (2012).

Early theories of morphological representations proposed architectures that could make the organization of the mental dictionary efficient in terms of information processing principles. The focus of these accounts was to achieve efficiency either in storage space or in processing cost. The *full-listing hypothesis* and *full- parsing* hypothesis represent two extremes on this debate.

The full-listing hypothesis (e.g., Butterworth, 1983), claims that regardless of their internal structure (either the base, inflected, derived, or compound forms) each word has its own representation. According to this hypothesis, words are stored without undergoing any decomposition processes and they are accessed as such. Therefore, the words *walk, walked, walkable, and crosswalk,* for example, would each have their own separate representations in the mental lexicon. Any possible relationship between words would emerge by product of (orthographic and phonological) form and meaning overlap, independent from their morphological relatedness. With direct access to each lexical entry, this type of organization offers processing efficiency. However, the amount of redundancy at the representational level is criticized as being unreasonable. For example, the words in languages with agglutinative morphology, such as Turkish, can frequently consist of 6-7 morphemes, with a possibility of more than tens of thousands of morphemes (Hankamer, 1989). In that regard, creating a separate entry for each form would result in massive redundancy in the storage of this framework.

The full-parsing hypothesis, on the other end of the debate, claims that only morphemic units are stored in memory. For lexical access to be successful, any complex word must first be parsed into its constituent morphemes. While this model brings an advantage to storage by eliminating the amount of redundancy proposed in the previous model, it severely increases the computational cost. The version proposed by Taft and Forster (1975) aimed to decrease this cost by reducing search space. This model was originally designed for prefixed derivations. According to the model, a decomposition process called *prefix stripping* was first applied to potentially complex words to separate the prefix from the base morpheme. The base, detached from its prefix, is then used as an access code to reach the lexical entry in the mental lexicon. When a matching lexical entry is found, the prefix and the base are reattached. As a result, the search space for the word *remind* would shrink from the number of words beginning with the prefix *re*- to the presumably smaller number of words beginning with *mi*. Yet, the model has been criticized for failing to deliver the promised efficiency, particularly in languages with a high number of pseudo-prefixed words, such as English and French (e.g., *re-* in *repertoire*). Failure of the model to correctly parse such words would increase the number of search steps, and eventually the processing cost (see Colé, Beauvillain, & Segui, 1989; Schreuder & Baayen, 1994, also Diependaele, Grainger, & Sandra, 2012, for a review)

Dual route models stay between these two extremes and combine features from both full-listing and full-parsing models (e.g., Caramazza, Laudanna, & Romani, 1988; Schreuder & Baayen, 1995). These models support both direct whole-word access and decomposition procedures simultaneously. However, how these routes operate varies depending on the assumptions made by a model: These routes can run in parallel, either in a cooperative or competitive fashion (e.g., Andrews, Miller, & Rayner, 2004; Kuperman, Schreuder, Bertram, & Baayen, 2009; Schreuder & Baayen, 1995), or only one route can operate at a time (e.g. Caramazza, Laudanna, & Romani, 1988). In the latter case, the criterion determining which route will be taken is determined by factors such as transparency, frequency, or familiarity of the words. For example, the Morphological Race Model (MRM) (Schreuder & Baayen, 1995) proposes a horse-race architecture in which both routes act in parallel and compete for lexical access: High frequency words take the direct route whereas low frequency words take the decomposition route. The Augmented Addressed Morphology model (AAM) (Caramazza, Laudanna, & Romani, 1988), on the other hand, prioritizes direct route access for familiar words and decomposition route access for unfamiliar words.

Among the models reviewed so far, those that advocated for a decomposition process for complex words proposed that morphemic units provide access to whole word representations. The Supralexical model (Giraudo & Grainger, 2000; Grainger, Colé, & Segui, 1991) takes a completely different approach, arguing that morphemic units are accessed after whole-word representations in the information processing hierarchy. This kind of organization, however, raised the question of how morphologically structured nonwords (e.g., *shootment*) can be identified within this framework since the whole word representation would never be accessed (as reviewed in Amenta & Crepaldi, 2012).

Thus far, the architecture of the models examined includes discrete representations of constituent morphemes and/or whole-words. The next two classes of models are amorphous approaches that reject any discrete lexical representations in the mental lexicon. The first class of models are connectionist models (e.g., Plaut & Gonnerman, 2000; Rueckl & Raveh, 1999; see also Hay & Baayen, 2005). This class of models propose a multilayer network comprised of units representing semantic and formal information as well as hidden units that adjust the strength of connections between these units. According to this view, morphology-like effects emerge as a consequence of distributed patterns of activation across these units.

The second class of models is Naive Discriminative Learning Model (NDL, Baayen et al. 2011) which has a simpler architecture with only a two-layer-network. Non-decompositional units like letter bigrams (e.g., *#w*, *wa*, *al*, *lk*, *k#*)

are represented in the input units and word forms (e.g. *walk*) are represented in the output units. According to this learning model, morphology-like effects emerge as a consequence of form and meaning co-occurrence: Frequent co-occurrence of input and output strengthens their connection and, as a result, leads to learning. Situations in which one of them is present while the other is not, on the other hand, weakens the strength of their connection and eventually results in unlearning. It should be noted that even in the absence of a discrete lexical representation, the model successfully replicated well-known morphological effects (Baayen et al., 2011).

To summarise, existing theoretical accounts provide a full spectrum of all possible scenarios for whether morphological effects are the result of explicit morphemic representations or the product of form and meaning co-occurrence, as well as whether a word's meaning is accessed directly, through distributional cues, or before or after morpho-orthographic segmentation of a complex word.

Repetition priming manipulations, where identical or similar stimuli are presented repeatedly, accounts for a considerable portion of the findings in the morphological processing literature. This technique is concerned with how the initial exposure to a word (prime) influences the visual processing of any given word (target) that is subsequently presented. The faciliatory effect is thought to emerge when the processing of the target benefits from the residual activation remaining after a prime has been processed (see Diependaele, Grainger, & Sandra, 2012). An increase in the processing ease of the target word is interpreted as evidence that the prime-target relationship is effective in visual word processing (for a description of variants of the priming task see Schmidtke & Kuperman, 2020).

Studies employing a repetition priming paradigm in various isolated word recognition tasks (e.g., lexical decision, naming, and fragment completion tasks) demonstrated convincingly that morphological structure plays a facilitatory role in visual word processing (e.g., Rueckl, 1990; Rueckl, Mikolinski, Raveh, Miner, Mars, 1997; Stanners, Neiser, Hernon, & Hall, 1979). This effect has been robustly observed across languages, within same modality as well as cross-modally (one visual, one auditory), regardless of the presentation duration of the prime (briefly or overt), the distance between prime and target (short, target immediately follows the prime; or long, there are intervening items between prime and target), or the type of complex word (e.g., compound, inflected, or derived words).

Although the evidence from isolated word recognition with repetition priming tasks draws a consistent picture, the nature of isolated word recognition tasks raises the question of whether morphological priming effect observed in single word paradigms can be generalized to natural text reading. Chapter 2 brings this question to readers' attention. Chapter 2 discusses the differences in experimental paradigms (i.e., isolated word recognition, isolated sentence reading, and natural text reading tasks) and how different experimental paradigms have produced dissimilar results in the field of reading in general (e.g., Kuperman et al., 2013), and in the field of morphological processing in particular (see discussion in Mousikou & Schroeder, 2019).

Chapter 2 focuses on a specific variant of the priming task, long-term morphological priming, in which the prime and target are separated by a range of intervening items. The influence of this type of priming has been robustly observed in isolated word recognition tasks among L1 readers (e.g., Münte, Say, Clahsen, Schiltz, & Kutas, 1999; Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997; Raveh & Rueckl, 2000). On the other hand, whether inflection priming, in general, exists in L2 reading received contradictory answers (see Clahsen & Neubauer, 2010; Jacob, Fleischhauer, & Clahsen, 2013; Jacob et al., 2017; Kirkici & Clahsen, 2013; Neubauer & Clahsen, 2009; Silva & Clahsen, 2008 for no or reduced inflection priming in L2 participants as compared to L1 controls; but also see Basnight-Brown et al., 2007; Coughlin & Tremblay, 2014; De Grauwe et al., 2014; Diependaele et al., 2011; Feldman et al., 2010; Foote, 2015 for similar results in both groups; also see Jacob, 2018 for a review). Furthermore, some argue that the reported inflection priming effects in L2 are due to orthographic overlap between morphologically related primes and targets, claiming that the L2 is more likely to rely on the orthographic surface form rather than the morphological structure (e.g., Diependaele et al., 2011).

To our knowledge, only one study (Kamienkowski et al., 2018) investigated long term priming effect in long text reading by analyzing an eye-tracking corpus of Spanish passages and reported no evidence of long-term priming. Therefore, Chapter 2 investigates whether Kamienkowski et al.'s conclusion is generalizable to other studies, and also languages. Chapter 2 makes use of the GECO database (Cop et al., 2017) of eye-tracking data recorded during book reading in English (by L1 and L2 readers) and Dutch (by L1 readers). These two languages, along with Spanish from Kamienkowski et al. (2018), created a continuum in terms of the opaqueness-transparency aspects of the languages examined in this paradigm: Spanish, has the most transparent orthography, followed by Dutch, while English has the most opaque orthography. As well, Chapter 2 addresses a gap in the literature regarding L1-L2 comparisons in long-term inflection priming, which was only investigated by few studies (e.g. De Grauwe, Lemhöfer, Willems, & Schriefers, 2014). Finally, Chapter 2 offers a systematic comparison of priming effect by analyzing all pairwise combinations of base-inflected forms as prime and target sets.

Contrary to the common finding of morphological priming in isolated word recognition tasks, Chapter 2 shows rather a controversial result in natural reading, reporting null effects in L1 and little evidence in L2, with a facilitatory effect of the inflected form for the base form (e.g. *trees – tree*). Chapter 2 calls for a greater theoretical attention to L1-L2 comparisons. The null findings in L1 reading inform revision of theoretical accounts to integrate natural reading behaviour, which is not limited to laboratory tasks' demands and artificial constraints. Additionally, this finding highlights the overall significance of studying natural reading behaviour in psycholinguistic research.

Chapters 3 and 4 turn our attention to the second lexical activity explored in this thesis: novel word learning. We will first detail the concept of word knowledge that will be common to these two sections, and then we will go over each objective in detail.

Novel word learning

Word knowledge is a multidimensional construct; knowing a word entails building knowledge in several facets of the word. According to Carter (2012), this entails the following seven abilities: (1) the ability to recognize a word in context and to use it properly, (2) the ability to predict how likely one is to encounter a word in a given context (e.g., the word *significant* is very likely to appear in an empirical research article), (3) the ability to understand the syntactic restrictions that determine the proper use of a word (e.g., a plural inflection following "those", *those feet*), and, the ability to recognize a word's internal structure and to produce acceptable words from a given word using word-formation rules (e.g., *redevelop* is derived from *develop, undeveloped* can be derived from *develop*) (4) the ability to understand a word's link to other words on the syntagmatic level, which includes associations between words from different word classes (e.g., *dog barks* or *dog attacks*), and on the paradigmatic level, which includes associations between words of the same word class (e.g., dog, pet, animal), (5) the ability to recognize the pragmatic and stylistic grounds for a word's selection (e.g., *would you mind* vs *can you*), (6) a solid understanding of word collocation patterns (e.g., *changing/flexible/shifting patterns*), (7) recognizing and using the fixed expressions in which a word appears (e.g., *it's been a long time since I've seen you; as far as I know; long time, no see*). Notably, Nation (2001) classifies this list into three groups: knowing a word involves knowing its form, (e.g., how it is written and pronounced, any constituent morphemes forming the word, awareness of its morphemic structure), its meaning (e.g., the link between form and meaning, its association with other words), and its use (e.g., common collocations, and contextual use). Fully mastering a word requires establishing receptive and productive knowledge in each of these components.

Word learning is a gradual process. Even for L1 speakers, obtaining full knowledge in all these facets of words may not be attainable for every word in their mental lexicon, especially for low frequency words (Schmitt, & Meara, 1997). Therefore, some aspects of word knowledge that develop with more experience, such as knowledge of the collocations in which the word appears, may reveal deeper insight into how well a person knows the word (e.g., *fair skin* but not *fair brown*) (Schmitt, 2010). Form learning and establishing the link between form and meaning, are, on the other hand, the very first components of word knowledge that can emerge after a few encounters to a novel word. As a result, these components would be suitable for assessing the vocabulary learning effectiveness of training delivered in a short period of time. To summarize, we recognize that word learning is a gradual process that occurs as a result of learning about multiple components of a word. Therefore, we would like to clarify that when we refer to novel word learning in Chapters 3 and 4, we are referring to the orthographic and semantic components of word knowledge, as measured by the post-tests of orthographic recognition and semantic recall and recognition performance.

Word knowledge is essential for both L1 and L2 speakers of a language to function in that language, and it may also be the most challenging component of learning a language (Nation, 2001; Saragi, Nation, & Meister, 1978; Schmitt, 2008). Both groups of learners carry the burden of selecting useful linguistic units from the context of discourse and of establishing the link between their form and meaning in the mental lexicon (Singleton, 1999). Furthermore, both groups may have the advantage (or disadvantage) of situations arising from individual differences: Individual differences such as a person's socioeconomic class, vocabulary development, and parents' education level can affect both L1 and L2 competence (Skehan, 1989). Similarly, L2 vocabulary size was found to be correlated with L1 vocabulary size, implying that those who are skilled at learning vocabulary in their first language are also good at learning vocabulary in their second (Henriksen, 2008).

In addition to these factors, certain characteristics of words facilitate vocabulary learning for both L1 and L2 speakers. The following are the characteristics of these words, according to Nation (1990): Learning a word would be easier if 1) a semantic representation already exists in the learner's lexicon (also Schmitt, 2010), 2) the form of a word provides obvious clues about its meaning (e.g., *buzz*, *burp*, *bizz*), 3) the meaning of a word is predictable from its form. For example, in "semantically transparent derived forms" such as *overachiever*, one can infer the meaning from the meanings of the constituent morphemes: *over-*, *achieve*, *-er* if they already have knowledge about the meanings of the constituents.

2. How does morphological priming affect novel word learning?

Chapter 3 focuses on base forms and their semantically transparent derived forms. The effort required to learn such words is highly correlated with one's morphology knowledge, as well as morphological awareness (the ability to recognize the internal structure of a word) abilities (see among many others, Carlisle, 2000; Mahony, Singson, & Mann, 2000). There are two potential learning mechanisms involved in inferring of the meaning of an unknown word based on pre-existing word knowledge and morphology knowledge: *generalization* and *decomposition.* Generalization refers to the ability to infer the (unfamiliar) meaning of a derived word (e.g. *farmer*) from the pre-existing knowledge of the base *(e.g., farm)* and of the meaning and function of the derivational morpheme (*-er*). Decomposition refers to the ability to infer the (unfamiliar) meaning of the base *(e.g., farm)* given the knowledge of the complex word *(e.g., farmer)* and the derivational morpheme (*e.g., -er*).

In word learning research, much focus has been given to understanding the generalization mechanism (from known base to unknown derived) (e.g. Anglin, 1993; Clark & Cohen, 1984; Clark & Hecht, 1982), while the decomposition mechanism (from known derived to unknown base) (e.g., Dawson et al., 2021; Ginestet, Shadbolt, Tucker, Bosse, & Deacon, 2021; Pacton, Foulin, Casalis, & Treiman, 2013) has received only little attention. Although base forms are more common and learned at a younger age, it is still possible for a learner to encounter a high frequency complex word first rather than its base (e.g., *computer* vs *compute*).

Chapter 3 investigates whether these two learning mechanisms, generalization and decomposition, produce equivalent learning outcomes or whether one is more efficient than the other, and if so, why. This question is examined in an incidental word learning paradigm with repetition priming, in which stories include all pairwise combinations of low frequency bases (*caltrop* "a spiked weapon against cavalry") and the novel derived forms of these bases (*caltroper* "a person who uses a spiked weapon against cavalry") as prime and target. Target words were preceded by the two repetitions of the prime. Each prime-target pair set exemplified either generalization (e.g., *caltrop-caltrop-CALTROPER*), decomposition (e.g., *caltroper-caltroper-CALTROP*), or control conditions (e.g., *caltrop-caltrop-CALTROP*, *caltroper-caltroper-CALTROPER*). The results demonstrate that decomposition is the more efficient mechanism for both orthographic and semantic learning than generalization, whereas the controls always led to optimal learning. Chapter 3 discusses reasons for the asymmetric effect of morphological structure on orthographic and semantic learning.

3. How do the well-known predictors of word learning – i.e., exposure and selective attention – interact in their effects on novel word learning. How many exposures to a novel word leads to a sufficient or to a maximum semantic or orthographic knowledge?

In chapter 4, we shift our attention from the influence of the internal structure of the repeated words to the influence of repetition itself, as well as the attention paid to these repetitions.

Repeated exposure to a novel word is regarded as a key factor for effective word learning (e.g., Ginestet, Valdois, Diard, & Bosse, 2020; Hulme, Barsky, & Rodd, 2019; Mohamed, 2018; Horst, 2005; Pellicer-Sánchez, 2016; Pellicer-Sánchez & Schmitt, 2010; Webb, 2008). Consequently, how exposure frequency influences word learning has been topic of many incidental word learning studies (Godfroid et al. 2017; Hulme et al. 2019; Mohamed 2018; Pellicer-Sánchez & Schmitt, 2010; Rott, 1999; Saragi, Nation, & Meister, 1978; Waring & Takaki , 2003; Teng, 2016). However, the number of exposures needed for word learning differs widely among these studies. Hulme et al. (2009), for example, reported that two encounters to a novel word would be sufficient for an L1 reader to learn the meaning of a novel word incidentally. Godfroid et al. (2017), on the other hand, reported that 8 to 10 encounters to a novel word would be expected for incidental learning to occur. This variation is even greater in L2 research, with reported values ranging from 6 (Rott, 1999) to 20 exposures (Waring & Takaki, 2003; also see Uchihara et al., 2019, for a review). Individual, word-level, or context-level variables account for some of this variance. However, a systematic comparison of L1 and L2 learners' word learning would help to understand how these two groups differ and what factors contribute to this difference by minimizing variability due to methodological differences. Despite the wide range of studies on novel word learning, only a few studies provide a systematic comparison of L1 and L2 adult learners on this issue (e.g., Cop, Dirix, Van Assche, Drieghe, & Duyck, 2017; Godfroid et al., 2017; Pellicer-Sánchez, Conklin, & Vilkaitė-Lozdienė, 2021; Pellicer-Sánchez 2016). As a result, the effects of word repetition on L1 and L2 learners' performance in incidental word learning is not yet to be fully understood.

Another gap in the literature is that the progressive nature of word learning is not fully addressed while examining the influence of repetition. Vocabulary learning is a gradual process. As a result, the number of exposures needed to gain *sufficient* and *maximum* word knowledge is likely to be different. However, most studies report the number of exposures that result in the highest performance in post-tests of learning (*maximum* word knowledge) while disregarding the minimum number of exposures that result in above-chance (*sufficient* word knowledge) performance in the tests (e.g, Hulme et al. 2019). It is even more difficult to determine these estimates for each component of word knowledge, such as orthographic or semantic knowledge, which do not necessarily develop at the same time or in the same way (e.g., Godfroid et al. 2017).

Another major predictor of word learning is considered to be selective attention to each occurrence of a novel word (Schmidt, 2010). Chapter 4 focuses on the preparation for action function of selective attention (see Bradley 2009, for a review). One line of word learning manipulation that taps into this function is intentional learning tasks, in which learners are prepared for the task with a pre-trial instruction that they will be tested on their word knowledge after the learning phase is complete. Intentional learning is expected to yield more learning gains than incidental learning, in which readers are not informed of upcoming tests (see review by Hulstijn, 2003). Apart from explicitly instructing participants they will be tested, we are not aware of any other experimental manipulation in the word learning literature that taps into this function of selective attention. Chapter 4 induces selective attention to occurrences of novel words by providing half of the participants with instructions about how many times a novel word was embedded in the upcoming paragraph. We believe that this manipulation will give readers the opportunity to strategically mobilize their attentional resources for the upcoming trial as an outcome of preparation action, resulting in improved word learning.

As briefly stated at the beginning of this section, eye-tracking can also provide insight into when the lexicalization of a word has started by examining the reading process of each individual occurrence. The use of eye tracking methods in word learning studies has grown in popularity in recent years. Eye-tracking studies examined how readers learned novel words in a variety of situations, such as when they were presented in isolation (Ginestet, Valdois, Diard, & Bosse, 2020), in a sentence (Chaffin et al., 2001; Joseph, Wonnacott, Forber, & Nation, 2014; Lowell & Morris, 2014), in short paragraphs (Pellicer-Sánchez 2016), or embedded in a novel (Godfroid et al., 2017). To our knowledge, only a few of them have compared the effects of repeated exposure to a novel word in an incidental reading context with both L1 and L2 readers (e.g. Godfroid et al., 2017; Pellicer-Sánchez 2016). It is worth noting that eye-tracking methodology can be particularly valuable to examine how each occurrence is processed and understanding how it contributes to learning; it is also a popular experimental paradigm for studying attention during reading (Rayner, 2009; Lowell & Morris, 2017).

To sum up, the first aim of Chapter 4 is to explore how the major predictors of novel word learning –i.e., attention and exposure– and interaction between them influence online and offline learning performance in L1 and L2 learners. The second aim of Chapter 4 is to investigate what level of exposure leads to *sufficient* knowledge (defined as minimal learning that will provide above chance level performance in the post tests) and to *maximum* knowledge (defined as the highest performance in post-tests) within the parameters of the study. We address these questions by implementing an incidental word learning paradigm, in which exposure was manipulated as the number of occurrences for a novel word in the passage (2, 4, or 8) and attention to novel words was manipulated by delivering attention-inducing instructions, while the control group received no instructions. L1 and L2 readers read passages while having their eye movements tracked and then were tested on the orthographic and semantic components of these novel words.

Chapter 4 shows that despite a language proficiency disadvantage, the performance of L2 learners in all tasks were comparable to those of the L1 learners. Crucially, L1 learners reached their optimal performance in the post-tests earlier (in the 4-exposure condition) while L2 learners' performance continued to improve up to the 8-exposure condition. We found little evidence for an effect of attention, with those who were given attention-inducing instructions showing a short-lived speed-up in eye-movements and faster recognition of novel words. Chapter 4 offers a comprehensive approach by bringing together two populations of novel word learners, online and offline tasks, and two major factors affecting word learning, attention and exposure.

Outline of the Thesis

Chapter 2 reports an eye-tracking corpus analysis with a focus on long-term inflection morphological priming in natural text reading in L1 English and Dutch, and L2 English. Chapter 3 reports an incidental word learning paradigm which examines how the morphological priming affects word learning in L1 adults. Chapter 3 compares the effectiveness of two learning mechanisms which are potentially used while inferring the (unfamiliar) meaning of a complex word from the pre-existing knowledge of the morphological constituents. Chapter 4 reports an incidental word learning paradigm using eye-tracking methodology. Chapter 4 investigates how exposure to and selective attention to a novel word interact in their effects on novel word learning in L1 and L2 speakers of English, and what level of exposure to a novel word leads to sufficient or to maximum semantic and orthographic knowledge. Finally, Chapter 5 summarizes and discusses the research findings.

References

Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology*, *3*, 232.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society of Research in Child Development, 58*(10, Serial No. 238).

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon*. John Wiley & Sons.

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology, 16*(1-2), 285-311.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*(3), 438.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*(1), 94-117.

Basnight-Brown, D., Chen, L., Hua, S., Kostić, A. and Feldman, L. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of Memory and Language*, 57(1), pp.65-80.

Butterworth, B. 1989. Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical Representation and Process.* Cambridge, MA: MIT Press.

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1–11.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant’s age. *Frontiers in Psychology*, 7, 1116.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*(3), 297-332.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing, 12*(3), 169-190.

Carter, R. (2012). *Vocabulary: Applied linguistic perspectives*. Routledge.

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: a study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 225.

Chomsky, N. (1965). Persistent topics in linguistic theory. *Diogenes*, *13*(51), 13-20.

Clahsen, H. and Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua*, 120(11), pp.2627-2637. http://dx.doi.org/10.1017/S1366728914000662

Clark, E. V., & Cohen, S. R. (1984). Productivity and memory for newly formed words. *Journal of Child Language*, *11*(3), 611-625.

Clark, E. V., & Hecht, B. F. (1982). Learning to coin agent and instrument nouns. *Cognition*, *12*(1), 1-24.

Coughlin, C. And Tremblay, A. (2014). Morphological decomposition in native and non-native French speakers. *Bilingualism: Language and Cognition*, *18*(3), pp. 524-542.

Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language*, *28*(1), 1-13.

Cop, U., Dirix, N., Van Assche, E., Drieghe, D., & Duyck, W. (2017). Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism: Language and Cognition*, *20*(4), 747-769.

Dawson, N., Rastle, K., & Ricketts, J., 2021. Bridging form and meaning: support from derivational suffixes in word learning. *Journal of Research in Reading*, *44*(1), pp.27-50.

De Grauwe, S., Willems, R. M., Rueschemeyer, S.-A., Lemhöfer, K., and Schriefers, H. (2014). Embodied language in first- and second-language speakers: neural correlates of processing motor verbs. *Neuropsychologia* 56, 334–349. doi: 10.1016/j.neuropsychologia.2014.02.00

Diependaele, K., Grainger, J., & Sandra, D. (2012). Derivational morphology and skilled reading. *Cambridge Handbook of Psycholinguistics*, 311-332.

Diependaele, K., Duñabeitia, J., Morris, J. and Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, 64(4), pp.344-358.

Dóczi, B. (2019). An overview of conceptual models and theories of lexical representation in the mental lexicon. *The Routledge Handbook of Vocabulary Studies*, 46-65.

Feldman, L. B., Kostić, A., Basnight-Brown, D. M., Durđević, D. F., & Pastizzo, M. J. (2010). Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism (Cambridge, England)*, *13*, 119–135. https://doi.org/10.1017/S1366728909990459

Foote, R. (2015). The Storage And Processing Of Morphologically Complex Words In L2 Spanish. *Studies in Second Language Acquisition*, 39(4), 735-767.

Frauenfelder, U. H., & Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In *Yearbook of morphology 1991* (pp. 165-183). Springer, Dordrecht.

Ginestet, E., Shadbolt, J., Tucker, R., Bosse, M. L., & Deacon, S. H. (2021). Orthographic learning and transfer of complex words: insights from eye tracking during reading and learning tasks. *Journal of Research in Reading*, *44*(1), 51-69.

Ginestet, E., Valdois, S., Diard, J., & Bosse, M. L. (2020). Orthographic learning of novel words in adults: effects of exposure and visual attention on eye movements. *Journal of Cognitive Psychology*, 1-20.

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review*, *8*(1), 127-131.

Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes*, *15*(4-5), 421-444.

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., Yoon, H. (2017). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition, 21*(3), 1–22. https://doi:10.1017/S1366728917000219

Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, *30*(3), 370-384.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), 662.

Hankamer, J. (1989). Morphological Parsing and. *Lexical representation and process*, 392.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review*, *106*(3), 491.

Hay, J. B., & Baayen, R. H. (2005). Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences*, *9*(7), 342-348.

Henriksen, B. (2008). Declarative lexical knowledge. In *Vocabulary and writing in a first and second language* (pp. 22-66). Palgrave Macmillan, London.

Hulme, R. C., Barsky, D., & Rodd, J. M. (2019). Incidental learning and long-term retention of new word meanings from stories: The effect of number of exposures. *Language Learning*, *69*(1), 18-43.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty, & M. H. Long (Eds.), The handbook of second language acquisition (pp. 349–381). (Blackwell handbooks in linguistics; No. 14). Malden, MA: Blackwell Publishing. https://doi.org/10.1002/9780470756492.ch12

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review, 61*(3), 355–382. https://doi.org/10.3138/cmlr.61.3.355

Jacob, G., Heyer, V. and Veríssimo, J. (2017). Aiming at the same target: A masked priming study directly comparing derivation and inflection in the second language. *International Journal of Bilingualism*, 22(6), 619-637.

Jacob, G., Fleischhauer, E. And Clahsen, H. (2013). Allomorphy and affixation in morphological processing: A cross-modal priming study with late bilinguals. *Bilingualism: Language and Cognition*, 16(4), 924-933.

Jarema, G., & Libben, G. (2007). *The mental lexicon: core perspectives*. Brill.

Joseph, H. S., Wonnacott, E., Forbes, P., & Nation, K. (2015). Corrigendum to "Becoming a written word: Eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading"[Cognition 133/1 238–248]. *Cognition*, *134*.

Kamienkowski, J. E., Carbajal, M. J., Bianchi, B., Sigman, M., & Shalom, D. E. (2018). Cumulative repetition effects across multiple readings of a word: evidence from eye movements. *Discourse Processes*, *55*(3), 256-271.

Kirkici, B., & Clahsen, H. (2013). Inflection and derivation in native and non-native language processing: Masked priming experiments on Turkish. *Bilingualism*, *16*(4), 776.

Kuperman, V., Jarema, G., & Libben, G. (2021). The Mental Lexicon Polylogue. *Polylogues on The Mental Lexicon: An exploration of fundamental issues and directions*, 1.

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *Quarterly Journal of Experimental Psychology*, *66*(3), 563-580.

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 876.

Libben, G. (2019). Words as action: Consequences for the monolingual and bilingual lexicon. Dans A. Tsedryk et C. Doe (dir.), *The description,* *measurement and pedagogy of words* (p. 14-33). Newcastle, Royaume-Uni : Cambridge Scholars Publishing.

Libben, G., & Schwieter, J. W. (2019). Lexical organization and reorganization in the multilingual mind. *The handbook of the Neuroscience of Multilingualism*, 297-312.

Libben, G., Goral, M., & Baayen, H. (2017). Morphological integration and the bilingual lexicon. *Bilingualism: A framework for understanding the mental lexicon*, *6*, 197.

Libben, M., Goral, M., & Libben, G. (Eds.). (2017). *Bilingualism: a framework for understanding the mental lexicon* (Vol. 6). John Benjamins Publishing Company.

Libben, G., & Jarema, G. (2002). Mental lexicon research in the new millennium. *Brain and language*, 81(1-3), 2-11.

Lowell, R., & Morris, R. K. (2014). Word length effects on novel words: Evidence from eye movements. *Attention, Perception, & Psychophysics*, *76*(1), 179-189.

Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and writing*, *12*(3), 191-218.

Mohamed, A. A. (2018). Exposure frequency in l2 reading: An eye-movement perspective on incidental vocabulary learning. *Studies in Second Language Acquisition*, 1–25. https://doi:10.1017/S0272263117000092

Mousikou, P., & Schroeder, S. (2019). Morphological processing in single-word and sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(5), 881.

Nation, I. S. (2001). *Learning vocabulary in another language*. Cambridge university press.

Nation, I.S.P. (1990). *Teaching & Learning Vocabulary.* Boston: Heinle & Heinle.

Neubauer, K., & Clahsen, H. (2009). Decomposition of inflected words in a second language: An experimental study of German participles. *Studies in Second Language Acquisition*, *31*(3), 403-435.

Pacton, S., Foulin, J., Casalis, S., & Treiman, R., 2013. Children benefit from morphological relatedness when they learn to spell new words. *Frontiers in Psychology*, 4.

Pellicer-Sánchez, A., Conklin, K., & Vilkaitė-Lozdienė, L. (2021). The effect of pre-reading instruction on vocabulary learning: An investigation of L1 and L2 readers' eye movements. *Language Learning*, *71*(1), 162-203.

Pellicer-Sánchez, A. (2016). Incidental l2 acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition*, *38*(1), 97–130. https://doi.org/10.1017/S0272263115000224

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a Foreign Language*, *22*(1), 31–55.

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing?. *Language and Cognitive Processes*, *15*(4-5), 445-485.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457-1506.

Rueckl, J. G., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language*, *68*(1-2), 110-117.

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C. S., & Mars, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language*, *36*(3), 382-405.

Rueckl, J. G. (1990). Similarity effects in word and pseudoword repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 374.

Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning*, *57*(2), 165–199. https://doi.org/10.1111/j.1467-9922.2007.00406.

Rott, S. (1999). The Effect Of Exposure Frequency On Intermediate Language Learners' incidental Vocabulary Acquisition And Retention Through Reading. *Studies In Second Language Acquisition*, *21*(4), 589-619.

Saragi, T., Nation, P., & Meister, G. (1978). Vocabulary learning and reading. *System*, 6, 72-80.

Schmidtke, D., & Kuperman, V. (2020). Psycholinguistic Methods and Tasks in Morphology. In *Oxford Research Encyclopedia of Linguistics*.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In *W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, Proceedings of CLaSIC 2010* (pp. 721–737). Singapore: National University of Singapore, Centre for Language Studies.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Springer.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329-363.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, *19*(1), 17-36.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. *Morphological Aspects of Language Processing*, *2*, 257-294.

Schreuder, R., & Baayen, R. H. (1994). Prefix stripping re-revisited. *Journal of Memory and Language*, *33*(3), 357-375.

Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, *4*(9), 353-361.

Silva, R. And Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11(2), 245-260.

Singleton, D. (1999). Lexis and the lexicon: some general considerations. *Exploring the Second Language Mental Lexicon*, 8-38.

Skehan, P. (1989). Language testing part II. *Language Teaching*, *22*(1), 1-13.

Stanners, R. F., Neiser, J. J., Hernon, W. P., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, *18*(4), 399-412.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A*, *57*(4), 745-765.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 638-647.

Teng, F. (2016). The effects of context and word exposure frequency on incidental vocabulary acquisition and retention through reading. *The Language Learning Journal*, *24*(2), 145–158.

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, *69*(3), 559-599.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader?.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*(1), 232–245.

Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging*, *18*(2), 332.

CHAPTER 2

Long-lag repetition priming in natural text reading: Little to no evidence for morphological effects.

This study has been submitted and currently under review in *The Mental Lexicon* as Coskun. M., Kuperman. V., Rueckl, J. (submitted). Long-lag repetition priming in natural text reading: Little to no evidence for morphological effects.

Abstract

Most of the empirical evidence that lays the ground for research on recognition of printed morphologically complex words comes from experimental paradigms employing morphological priming, e.g., exposure to morphologically related forms. Furthermore, most of these paradigms rely on context-less presentation of isolated words. We examined whether well-established morphological priming effects (i.e., faster recognition of a word preceded by a morphologically related word) are observable under more natural conditions of fluent text reading. Using the GECO database of eye-movements recorded during the reading of a novel, we examined the long-lag morphological and identity priming in one's first language (L1, English and Dutch) or second language (L2, English). While the effects of identity priming were ubiquitous, no evidence of morphological priming was observed in the L1 eye-movement record and little evidence was found in L2. We discuss implications of these findings for ecological validity and generalizability of select current theories of morphological processing.

Introduction

Repetition priming – the repeated presentation of identical or similar stimuli is a very popular type of experimentation in word recognition research. The main premise of a priming task that an initial exposure to a prime stimulus influences the subsequent processing of a target, and the degree and nature of the overlap between the prime and the target determine how much cognitive effort the recognition of the target would require. By comparing different types of priming – ranging from the maximum overlap as in the identity priming (prime *car*, target *car*) to the minimum overlap where the prime and the target are unrelated (*car* - *bog*) – researchers expect to pin down behavioral consequences of orthographic, phonological, morphological, semantic, and other relationships between words. Our main interest is in morphological priming, i.e., studies in which the prime and the target are lexical items sharing a morpheme, e.g., car – cars (for an early influential study see Stanners, Neiser, Hernon, & Hall, 1979), see details below. In many priming experiments, the task that enables measurement of the priming effect is lexical decision, and the quantity of interest is the change in speed or accuracy of identifying target words as a function of the relationship between the prime and the target. Some studies couple lexical decision with the registration of the neurophysiological brain activity: in these cases, recordings of electro-magnetic or hemodynamic signals complement the behavioral signatures of morphological processing (see review by Leminen, Smolka, Duñabeitia, & Pliatsikas, 2019).

The preponderance of lexical-decision-with-priming as an experimental paradigm of choice is high, so much so that several recent review papers and handbook chapters purporting to cover the entire field of morphological processing have explicitly confined themselves to outcomes of this paradigm (e.g., Amenta & Crepaldi, 2012; Milin, Smolka, & Feldman, 2017; but see Bertram, 2011; Marslen-Wilson, 2007; Schmidtke & Kuperman, 2020). The paradigm is diverse and ranges from the most popular and influential masked priming tasks with very short lags between primes and target to long-term priming, in which multiple stimuli intervene between the prime and the target. In the last few decades, lexical decision studies of morphological priming played a prominent role in supplying the evidence base for theoretical accounts of morphological processing (e.g., Crepaldi, Rastle, Coltheart, & Nickels, 2010; Forster, Davis, Schoknecht, & Carter, 1987; Frost, Forster, & Deutsch, 1997; Gonnerman, Seidenberg, & Andersen, 2007; Grainger, Colé, & Segui, 1991; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Rastle, Davis, Marslen-Wilson, & Tyler, 2000). This reliance of both the evidence base and the theoretical landscape of the field on the results of lexical decision studies of morphological priming begs a question of whether these results generalize over other experimental paradigms. This question is at the core of the present study.

We examine a specific type of morphological priming – a long-lag inflection priming occurring when the primes and the targets are represented either by base forms or regular inflected forms of English and Dutch nouns, verbs, and adjectives (e.g., argue/argues/argued/arguing; tree/trees) and are typically separated by multiple intervening words. The examination makes use of the eye-tracking record of reading a book-length text in either one's first (L1 English and Dutch) or second (L2 Dutch) language (Cop, Dirix, Drieghe, & Duyck, 2017). While this type of priming has been robustly observed in paradigms where words presented in isolation (lexical decision and fragment completion, see below), we capitalized on the high ecological validity of eye-tracking which offers a non-invasive and reliable measurement of naturally occurring reading behavior (Rayner, 1998). The study aims to establish (i) whether long-lag inflection priming occurs in natural reading, (ii) how the nature of the prime and the target (given as all combinations of base and inflected word forms) affects the priming effects, and (iii) what aspects of the eye-movement behavior, if any, are influenced by priming and what cognitive processes are likely to give rise to these effects. In the remainder of the Introduction, we provide a detailed description of the priming task in question, review the literature that compares experimental paradigms of word processing research, and outline the logic of this study.

Inflection priming in isolated word recognition

Morphological research manipulates what the prime and the target are (typically one being a simplex word like *talk* and the other a complex inflected form like *talks/talked* or a derived form like *talker*), and the degree of an orthographic, phonological, or semantic overlap between them. Other manipulated dimensions include, among many others, the modalities in which the prime and the target are presented; the duration of a participant's exposure to the prime; whether the prime is masked by visual or auditory noise; and how far apart the presentations of the prime and the target are in the experiment and what items intervene (for a description of variants of the priming task see e.g., Schmidtke & Kuperman, 2020). As indicated above, this study focuses on a specific variant of the priming task, i.e., long-term (long-lag) morphological priming. The choice of this type of priming is dictated by two considerations. First, unlike several other types of priming, repetition of same words or of morphological variants of the words with additional intervening unrelated words is part and parcel of naturally occurring texts: thus, long-lag priming can be studied comparatively in the eye-tracking record of text reading and isolated word recognition. Second, in English and Dutch regular inflections are maximally similar to the base forms, i.e., they show a very high degree of an orthographic and semantic overlap, typically only differing in one or two syntactic features (e.g., present vs past tense; singular vs plural number; or person). This overlap maximizes the likelihood of observing the morphological priming effect in our data.

Inflection priming has been robustly observed across languages and manipulations. For instance, two lexical decision experiments in English by Raveh and Rueckl (2000) reported significant priming effects when the inflected prime was separated from the simplex target (e.g., boiled-boil) by 8-13 intervening items (see also a lexical decision study with inflected and derived forms by Rueckl & Aichler, 2008; a fragment completion study with mixed inflected and derived stimuli by Rueckl & Galantuci, 2005 and with irregular inflections by Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997; and Stanners et al., 1979). Similar effects were found in regular inflections in English, French, German, and Spanish (Münte, Say, Clahsen, Schiltz, & Kutas, 1999; Rodriguez-Fornells, Münte, & Clahsen, 2002; Royle, Drury, Bourguignon, & Steinhauer, 2012; Weyerts, Münte, Smid, & Heinze, 1996), also in children (see a fragment completion study by Feldman et al., 2002). Long-term inflection priming was also reported in Hebrew and Serbo-Croatian, at short lags (0) and long lags (10), with an additional manipulation of alphabetic systems (Roman vs Cyrillic), and both when an inflected prime preceded a base-form target and when the base-form prime preceded an inflected target (Bentin & Feldman, 1990; Feldman, 1992, Feldman & Bentin, 1994; also see review by Feldman & Andjelković, 1992). Thus, long-term inflection priming exists in isolated word recognition among L1 readers.

There is less consensus regarding inflection priming in L2 readers. As reviewed in Jacob (2018), much relevant research reports no or reduced inflection priming in L2 participants as compared to L1 controls (e.g., Clahsen & Neubauer, 2010; Jacob, Fleischhauer, & Clahsen, 2013; Jacob, Heyer, & Veríssimo, 2017; Kirkici & Clahsen, 2013; Neubauer & Clahsen, 2009; Silva & Clahsen, 2008). This is despite robust effects of identity priming and either full or partial effects of derivation priming in L2 readers, which demonstrate that they are generally sensitive to manipulations in the overt or masked priming paradigms. The argument that L2 readers make less use of the word's morphological structure during recognition is not universally accepted. Some papers do find inflection priming effects in L2 word recognition comparable to those effects in L1 controls and argue for similar sensitivity of L1 and L2 readers to morphological structure (Basnight-Brown et al., 2007; Coughlin & Tremblay, 2014; De Grauwe et al., 2014; Diependaele et al., 2011; Feldman et al., 2010; Foote, 2015). A remaining question for proponents of inflection priming in L2 is whether these effects are truly morphological or rather reflect a greater reliance of L2 readers on surface orthographic forms and the orthographic overlap between morphologically related primes and targets. Evidence for this stems from studies reporting the effect of the orthographic priming only in the L2 group but not in the L1 group (Diependaele et al., 2011; Feldman et al., 2010, Heyer & Clahsen, 2015; Ciaccio & Jacobs, 2019). For example, Heyer and Clahsen (2015) compared orthographically and morphologically similar prime-target pairs (scanner-scan vs scandal-scan) in a masked priming study. While priming for morphologically similar words was seen both in native and non-native groups, facilitation effects for orthographically similar words were only seen in the non-native group.

Yet, the findings showing the L1 and L2 contrast in processing inflected words are not sufficient to conclude that L1 and L2 morphological processing systems rely on the distinct mechanisms. One group of studies suggest that the L1-L2 performances differences can be explained by cognitive limitations of the L2 group (e.g., limited working memory capacity, slower processing speed) or by L1 background influence on L2 processing (McDonald, 2006; see also Clahsen & Felser, 2006; Neubauer & Clahsen, 2009 for a review). For example, McDonald (2006) showed that factors like noise stress or high working memory demand made the L1 group perform in the same way as the L2 group in a speeded sentence grammaticality judgement test, in which grammaticality violations were caused by mismatched inflectional suffixes. On the other hand, other studies suggest that especially late learners of L2, who are less experienced in L1 rely more on lexical storage rather than combinatorial processing in order to process morphologically complex forms (e.g., Silva & Clahsen, 2008).

Thus, the presence and magnitude of long-term inflection priming varies across isolated-word experimental paradigms, regularity of stimuli, and language background. Below, we examine long-term inflection priming in both L1 and L2 readers.

Isolated word recognition vs sentence reading vs text reading

A convergence of evidence between paradigms is an important desideratum of any research field. In the case of morphological research where much evidence hinges on isolated word recognition the need is particularly strong, because this paradigm deviates from natural reading in a number of critical respects (Liversedge, Blythe, & Drieghe, 2012; Rayner & Liversedge, 2011). In this section, we discuss possible processing consequences of using isolated word recognition tasks in general, with or without priming. Unlike the texts that humans read, a stimulus list of a typical lexical decision experiment contains 50% or more of non-lexical items. Furthermore, the percentage of morphologically complex stimuli tends to be much greater in experiments on morphological processing than in natural language. Unlike a continuous reading of semantically meaningful and coherent texts for comprehension, all stimuli are presented in isolation for a meta-linguistic judgment that is hardly in demand outside of a laboratory and can be successfully done without recourse to at least some levels of linguistic representation (Grainger et al. 2012; Linke et al., 2017). Some priming manipulations done in the lexical decision framework also tax the participant's perceptual system by limiting the duration of their exposure to stimuli and introducing noise.

As a result of these differences, some data harvested from lexical decision tasks only bear a limited resemblance to behavioral patterns found in tasks involving text reading. Comparisons by Kuperman et al. (2013) discovered low correlations between the lexical decision response times and fixation times to the same words in English and Dutch sentence reading. Shared variance was further reduced to a negligibly small amount when effects of word length and frequency were partialled out from lexical decision latencies and word-reading times.

Indeed, different experimental paradigms have produced dissimilar results in the field of morphological processing as well (see discussion in Mousikou & Schroeder, 2019). To give a few examples, Masson and MacLeod (1992) demonstrated that the magnitude of morphological priming observed in isolated word recognition is drastically reduced when same words are presented in context and require semantic integration. Schmidtke et al.'s (2020) analyses of a corpus of English eye-tracking sentence-reading data indicated a number of morphological effects on compound recognition that have been reported in the lexical decision literature but are not detected even in a highly-powered eye-movement record (see also Juhasz et al., 2003). Schmidtke et al. (2017) and Schmidtke and Kuperman (2019) reported discrepancies in the estimated time-course of processing derived words and compounds, respectively, in a comparison of eye-movements recorded during sentence reading and in lexical decision experiments and mega-studies. Taken together, these discrepancies suggest that many research outcomes based on the lexical decision paradigm may not readily translate to the case of sentence reading or text reading (see review by Bertram, 2011; but see also Mousikou & Schroeder, 2019; Marrelli & Luzzatti, 2012, for parallel results between reading in isolation vs in sentence context).

Eye-tracking studies of morphological priming

We now turn to the eye-tracking studies of morphological priming. The study by Paterson et al. (2011) included primes and targets in carrier sentences in English, with a few intervening words, and manipulated the degree of overlap from a semantically transparent derivation (marshy – marsh), to opaque derivation (secretary – secret) to morphologically unrelated but orthographically related case (extract - extra). Paterson et al. observed a semantically mediated priming both in the early and late eye-movement measures. Furthermore, Paterson et al. found that in all three conditions readers skipped the target words more frequently and made fewer regressions when targets followed primes rather than controls. As all three conditions had orthographic overlap at the onset of prime and target, the observed advantage in skip rate was attributed to parafoveal pre-processing. Priming was also reported by Mousikou and Schrouder's (2019) study in German. They administered three lexical decision and one sentence-reading eye-tracking experiment with affixed (prefixed and suffixed) primes and morphologically simple targets. While their lexical decision studies used masked priming (the prime was shown for 50 ms and then masked by hashtag symbols), the eye-tracking used a clever fast priming manipulation in which the prime was masked from the reader's parafovea until the eyes of the reader crossed the invisible boundary before landing on the prime; the prime was then shown for 50 ms and finally replaced in the same area on the screen by the target word. The lexical decision and eye-tracking data revealed the same patterns, indicating early processing of stems embedded in the affixed words.

Morphological priming from the parafoveal primes was also reported by a recent study by Dann, Veldre, and Andrews (2021). Dann et al. employed the gaze-contingent boundary paradigm (Rayner, 1975) in sentence reading. Unlike the fast priming paradigm, the prime, also called preview, is changed to the target word as soon as a reader's eye crosses the invisible boundary. A reader is usually unaware of this change because of saccadic suppression (Matin, 1974). The targets were either suffixed (e.g., stressful) or prefixed words (e.g., mistrust). The previews were either identical (e.g., stressful), morphologically decomposable non-words (e.g., stressary), or non-words containing a non-morphological suffix (e.g., stressard). They found the same durations for suffixed words (e.g., stressful) in case of an identical parafoveal preview in comparison to a morphologically decomposable non-word preview (e.g., stressary), but slower durations when the preview contained a non- morphological suffix (e.g., stressard), indicating that the morphological structure of a word can be obtained from parafoveal vision before the word is fixated. However, the influence of morphological structure on the early stages of word recognition is only limited to suffixed words but not prefixed words. The eye-tracking results revealed that the suffixed previews provided the same benefit as the identity previews in early fixations, with shorter first fixation duration and single fixation durations to the target words. In sum, studies of morphological derivation observed both early and late effects of semantically transparent primes and varied in their conclusions regarding the role of the primes that were semantically opaque and morphologically unrelated.

While sentence reading studies obviate the drawbacks of isolated word recognition, their ecological validity is limited by the fact that sentences are carefully constructed to (over)represent stimuli of interest and are unrelated to other sentences in the stimulus list. Indeed, differences have been found within the eye-tracking paradigm between the reading of isolated sentences and that of long stories or novels. Dirix et al. (2019) demonstrated relatively low correlations between response times to the same words in lexical decision, sentence reading, and reading of a novel, as well as a sizable variability in the magnitude of word length and frequency effects on lexical decision response times and durational eye-movement measures. Radach et al. (2008) also observed shorter reading times to the words if they were read as part of a long story rather than a sentence. Thus, the question remains whether morphological priming is found in less constrained natural reading.

One study that addressed this question directly answered this question in the negative. Kamienkowski et al.'s (2018) analyzed an eye-tracking corpus of Spanish passage reading with a focus on the effect of identity priming (word repetition, tree-tree) and inflection priming (base form - inflected form, tree-trees). Specifically, they considered the first two occurrences of the word forms (using the base form as a baseline for the identity or morphological priming) and examined whether indeed the second target was read faster, while taking into account the number of intervening words between the two occurrences of interest, their frequency and other relevant controls. Kamienkowski et al.'s (2018) observed a healthy effect of identity priming: a word form – either a base form or an inflected form – seen the second time in the text was processed faster than it occurred the first time around. Critically, they found no indication of morphological priming: neither a word form following its base form nor a base form occurring after its inflected form elicited faster processing times compared to those that were not preceded by a morphological relative. This null effect indicates that long-term inflection priming, robustly reported across languages and levels of proficiency in the lexical decision paradigm, does not find support in Spanish proficient readers in a situation that more closely resembles natural reading.

What may be causing these null effects in natural reading? One explanation could be that, contrary to single word processing, natural reading is not conducive to the emergence of morphological priming. Reading a word in isolation versus in the context of a complex text may tap into overlapping but different component skills and cognitive operations and elicit different processing. This argument is made, inter alia, in the Multiple Read-Out Model of visual word recognition (Grainger & Jacob, 1996) which highlights how task demands influence the underlying mechanisms that play a role in the visual word identification process. For instance, when required to pass a meta-linguistic judgment as soon as possible in a lexical decision task, guessing is likely to be involved. The process of guessing may partly explain, on this theoretical account, how the implicit identification of a (pseudo)root in the prime word (e.g., *corn-er*, *farm-er*) in the masked priming paradigm may be sufficient to increase the activation in all the lexical nodes that share such root (e.g., *corn*, *farm*), leading to a priming effect. On the contrary, the account claims that during text reading, the reader must identify the target word exactly, so the previous exposure to a morphologically related word is not enough to exert a priming effect.

Furthermore, a single word recognition task does not provide a reader with any additional information. Thus, the prime becomes the primary source of information that activates the related words, which is another reason explaining robust morphological priming effects being likely to be observed in this task. In natural reading, on the other hand, morphosyntactic, semantic and other contextual cues narrow down target words to be recognized. Furthermore, text reading affords a parafoveal preview of the word, which altogether make the processing of a word easier. Thus, the benefit of morphological priming may not arise in natural reading especially when there is a long distance between the prime and the target. This null effect is particularly expected when the readers' proficiency is high, like in the case of healthy adults reading a text, written in their own language. In this case, the automatized activation of the whole word is so fast that any morphological priming effect may be obscured by the speed of whole word identification. In such a case, eye-movements are affected only by the psycholinguistic features (e.g., word frequency, root frequency, length in letters, etc.) of the target itself. On the other hand, when reading mastery is not reached, as in the case of L2 readers or young readers with dyslexia, a role of morphological priming might be expected, as the whole-word activation is not yet automatized (see Marelli, Traficante & Burani, 2020 for a review; Burani et al. 2008; Burani 2010; Carlisle & Stone 2005; Deacon, Tong, & Mimeau 2019; Mann & Singson, 2003; Traficante et al. 2011).This discrepancy between the paradigms forms a basis of the present study[1].

Since identity priming (where the prime and the target are identical stimuli) is frequently used in morphological research as a baseline against which the priming effect of other conditions is evaluated, we also point to its well-examined equivalent in eye-tracking research of reading, i.e., the word repetition effect. Multiple readings of the same text lead to decreased reading times and higher skipping rates for words that have been already seen in the previous readings (Chamberland et al., 2013; Raney & Rayner, 1995). Moreover, the reduction of the processing effort due to repetition was stronger in low-frequency than in high-frequency words (Rayner, Raney, & Pollatsek, 1995). The same pattern of effects has been observed both in L1 and L2 readers, both for familiar and unfamiliar words, when those words occurred multiple times in the same text (see among others Godfroid et al., 2018; Kamienkowski et al., 2018). We expect to find identity priming, or the word repetition effect, in our study as well.

The present study

As indicated above, the main motivation for this study in long-lag inflection priming originates from the need for validating results obtained within paradigms using isolated word recognition against other, arguably more ecologically valid experimental paradigms. We use the Ghent Eye-Tracking Corpus GECO (Cop et al., 2017) of L1 reading in English and Dutch and L2 reading in English (see details below) to examine the presence and magnitude of inflection priming. The inclusion of these two languages (along with Spanish from Kamienkowski et al., 2018) create a continuum in terms of the opaqueness-transparency aspects of the languages examined in this paradigm. Among these languages, Spanish has the most transparent orthography, followed by Dutch, and English has the most opaque orthography. Additionally, this study expands the scope of Kamienkowski et al.'s study by examining both L1 and L2 reading. As a baseline comparison, we also look at identity priming, i.e., repetition of the same base form or inflected form. Specifically, we consider all types of pairwise combinations of base forms and inflected forms as examples of long-term identity and inflection priming in natural texts read for comprehension. If – in line with the lexical decision findings (reviewed above) and contra Spanish data of Kamienkowski et al.'s (2018) – we observe an effect of morphological priming on eye-movement measures of word recognition, the evidence base of morphological research will be enhanced by a cross-paradigm convergence. If no morphological priming is observed in the eye-tracking data, one needs to entertain a possibility that this morphological effect is due to specific task demands imposed by either the isolated word recognition, the nature of the stimulus list or other artifactual dimensions of a particular research paradigm.

Methods

Materials

To investigate how identity and morphological priming influences reading of this word and morphologically related forms, we used the Ghent Eye-Tracking Corpus, GECO, (Cop, Dirix, Drieghe & Duyck, 2017). This corpus contains eye movement measures from 14 English monolinguals (age range: 18 – 36, *M* = 21.8, *SD* = 5.6) reading a novel in English and 19[1] Dutch– English bilinguals (age range: 18 – 24; *M* = 21.2, *SD* = 2.2) reading the same novel both in their first language (L1, Dutch) and in their second language (L2, English). The English original and the Dutch translation of the novel *The Mysterious Affair at Styles* by Agatha Christie (1920) were read during four sessions of one hour each on the same day. The monolingual group read the entire book in English. The bilingual group read half of the book in their first language and half in English. The order of the languages was counterbalanced such that half of the group read the first part of the book in Dutch and the last part in English whereas the other half did the reverse. The majority of L2 readers were reported in Cop et al. (2017) as being upper intermediate and advanced speakers of English based on the norms of the

[1] One of the bilinguals only read the half of the book in English.

LexTALE test (Lemhöfer & Broersma, 2012). Dutch and English native speakers were also reported as being equally proficient in their first languages.

The corpus consisted of 59,716 Dutch (5,575 unique types) and 54,364 English (5,012 unique types) words. Content words represented 41.9% of the Dutch data (of those, 31.6% were nouns, 52.3% were verbs, 16.1% were adjectives). The length of those words ranged from 2 to 22 (*M* = 6.07, *SD* = 2.69). Content words represented 43.9% of the English data (of those, 39.1% were nouns, 50.1% were verbs, 10.7% were adjectives). The length of English content words ranged from 2 to 17 (*M* = 5.37, *SD* = 2.28). For further details of the corpus, see Cop et al. (2017).

We chose to use the GECO since its texts are long and lexically diverse enough to contain an adequate number of repeated words and morphological forms and thus enable an examination of the effect of word priming. Also, the public availability of eye tracking measures from monolingual and bilingual readers of GECO enabled an easy comparison of different languages and of reading patterns of native versus non-native readers.

Variables

Dependent variables. GECO data files make available a range of eye-tracking measures, each potentially reflective of a different (aspect of) cognitive process implicated in word and text reading. We examined both early eye-tracking measures including first fixation duration, gaze duration, and skipping rate, and also late eye-tracking measures including total fixation time and regression rates (Rayner, 2009). Early measures are generally associated with the initial stages of word processing such as word recognition and lexical access (Conklin & Pellicer-Sánchez, 2016) whereas late measures tap into the later stages of processing, which may reflect an effort to correct a misreading, to resolve an ambiguity, or to integrate semantic knowledge (Libben & Titone, 2009). More specifically, we considered five measures as dependent variables including duration of the first fixation on the word (an index of the effort of initial lexical access and word decoding), gaze duration (i.e., the summed duration of all fixations landing on the word before the eyes leave the word for the first time and an early index of word recognition effort), and total fixation time (i.e., the summed duration of all fixations landing on the word and a cumulative index of word processing effort including semantic integration with context; Boston et al. 2008). Other measures included skipping rate (the likelihood of not fixating on the word even once, an index of the ease of recognizing the word while it is available in the parafoveal preview) and regression rate (the likelihood of a look-back to the word after the eyes have moved forward from the word).

Independent variables. An independent variable of critical interest was the type of priming that repetition of words or their related forms in the read texts gave rise to. To define this variable, we distinguished between a base form of target words (e.g., a dictionary form like *tree; advise*) and an inflected form of those words (e.g., *trees; advising, advised, advises*). All (base and inflected) forms of the same word can be attributed to the same lemma which, in English and Dutch, is typically labeled as a base form (*tree*; *advise*). Within the specified criteria, the 1242 words appeared more than once in both English and Dutch versions of the novel (for the identity condition). The majority of them are only repeated twice: In the English edition, this accounted for 35% (434 words) of the repeated words, while in the Dutch edition, it accounted for 25% (312 words). These rates drop dramatically as the number of repetitions increases. In the English version of the data, 16% of words were repeated three times, while 10% were repeated four times. For the Dutch version, these numbers were 17 % and 11%, respectively. We projected that focusing on words that were repeated more than twice in the analysis would drastically reduce the number of data points, lowering the statistical power of our analyses. Therefore, the analysis below considered the first two occurrences of either the base form or the inflected word form in the novel. Each occurrence of the base form and inflected form was coded for its morphological type (*bf* for base form and *if* for inflected form) and order of occurrence (1 or 2). We considered two sets of morphological forms separately. One consisted of (i) inflected forms in the first position (unprimed by any other morphological form of that word at any lag, e.g., *trees*), (ii) inflected forms in the second position, preceded by the same identical forms (e.g., trees – *trees*), and (iii) inflected forms in the second position, preceded by the respective base form (e.g., tree – *trees*). Controlling statistically for the influence of other covariates and between-participants and -items variation (see below), a regression model fitted to lexical items in (i), (ii) and (iii) estimated both identity priming – the difference between (i) and (ii) – and morphological priming – the difference between (i) and (iii) – for inflected forms as targets. The second set consisted of a similarly organized separate model that estimated morphological (trees – tree) and identity (tree – tree) priming for base form targets.

Control variables. The ordinal position of a word in the text was reported to be a significant predictor of word reading time in the regression models reported by Kamienkowski et al. (2018), see also Kuperman, Matsuki and Van Dyke (2018). Thus, the position from the beginning of the novel that a target word instance has in the text was included as a covariate in our models.

Participants read the novel in 2-4 sections. We identified cases of priming for the entire novel and within specific sections. Both sample sizes and observed priming effects were very similar whether we took the within-section constraint into account or not. Below we report analyses based on the entire novel. We considered the distance between the two wordform occurrences as a difference in their ordinal numbers in the text.

The distance between the selected words ranged between 1 and 50,087 words (M = 9,749, SD = 11,448) in English L1 material. This distance ranged between 1 and 27,380 words (M = 4,935, SD = 5702) in English L2 material and it ranged between 1 and 30,491 words (M = 5484, SD = 6180) in Dutch L1 material. The vast majority of the prime-target pairs in all corpora and conditions were separated by more than one hundred of words (e.g., the median lag in the English texts was 6,357 and 93% of lags were over 100 words), residing in different sentences and often chapters. This very long lag may differ from the commonly used experimental settings but is reflective of the naturally occurring distributional patterns of language use. We explored several operationalizations of lag, ranging from the continuous metric of distance to the binning into discrete categories. None of the metrics of the distance affected any of the results below (similar to Kamienkowski et al., 2018), and we did not include lag as a variable it in the reported models below.

Word frequency and word length were also included as covariates into the models. The frequency counts of the words were extracted from SUBTLEX-US (Brysbaert, New, & Keuleers, 2012) and SUBTLEX-NL (Keuleers, Brysbaert, & New, 2010) databases for English and Dutch parts of the GECO, respectively. Frequency counts were log-transformed to reduce skewness in the distributions and the influence of the outliers. We made use of the word class and lemma annotations of the English and Dutch words in SUBTLEX-US and SUBTLEX-NL corpora respectively for identifying base forms and inflected forms of the target words. The word class categories of the words were also included into the model as a control variable: we only considered verbs, adjectives, and nouns.

Statistical considerations

We modeled priming effects separately for each of the three subcorpora of GECO: English L1, Dutch L1, and English L2. Statistical data analyses were conducted in R version 3.6.3 (R Core Team, 2020). We applied linear mixed-effects regression models (using *lmer* function from the *lme4* package, Bates, Maechler, Bolker, & Walker, 2015) to account for variability between items and participants. The *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017) was used to estimate *p*-values for fixed-effects with Satterthwaite's approximation for degrees of freedom. Durational eye-movement measures were modelled using the Gaussian family, while binary skipping and regression indices using logistic regression. We used the *bobyqa* optimizer to minimize possible convergence issues.

A total of 30 regression models (5 dependent variables x 2 types of target (base and inflected form) x 3 subcorpora) were fitted to the data. All models included control variables presented above and random intercepts by participant and word. Inclusion of random slopes for the morphological type of the target led to convergence errors and was not implemented.

Because multiple dependent variables were studied, we additionally applied the Bonferroni correction to the *p*-values estimated for the five dependent variables in each subcorpus (von der Malsburg & Angele, 2017). Estimation of statistical power based on generalized linear mixed effects models, reported below, used R package *simr* version 1.0.5 (Green & MacLeod, 2016).

Results

The original datasets contained 774,015 observations for the English L1 subcorpus, 549,290 for the Dutch L1 subcorpus, and 534,154 for the English L2 subcorpus. We excluded English and Dutch words that are not included in the frequency lists of the SUBTLEX-US and SUBTLEX-NL corpora, respectively. We confined our consideration to verbs, adjectives, and nouns only. Furthermore, we only included lemmas with regularly inflected word forms (e.g., tree/trees but not child/children). We limit our data set to the first two occurrences of a word, either an inflected or base form: all subsequent occurrences of the word in any form were dropped from consideration. We then coded words for skipping and conducted the analyses of skips based on this dataset. We further removed words with durations shorter than 80 ms and longer than the 99th percentile of the total fixation duration range. The summary of the eye movement characteristics and resulting sample sizes (i.e., the number of unique words) are presented in Table 1.

As described in the Methods, we fitted a total of 30 models to subsets of data associated with different eye-movement dependent variables, different target word types (base form or inflected form) and subcorpora (English L1, Dutch L1, and English L2). Regression models and conditional means are reported in Appendix, Table S1-S5.

Table 2 presents the results of the planned comparisons representing either morphological priming or identity priming. Statistically significant findings (after correcting for multiple comparisons) are shown in bold. We discuss the two types of priming in turn.

Morphological priming

The only evidence in favor of morphological priming was observed in L2 reading of the English text by Dutch students. Specifically, we found morphological priming when an inflected form was a prime and base form the target (if – BF, trees - tree). An earlier occurrence of the inflected word sped up the total fixation time to the base form by around 15 ms, compared to the unprimed base form. This morphological priming effect was also observed in gaze durations, with a 10 ms reduction of reading time. When corrected for multiple comparisons within each group of five dependent variables, no other instance of morphological priming reached significance. This absence of effect held true for English L1, Dutch L1 and a slate of oculomotor measures representing the entire time-course of word recognition.

Table 1. Summary of eye-movement characteristics in the three subcorpora. BF – base forms in the first position; bf_BF – base forms preceded by base forms. bf_IF – inflected forms preceded by base forms. IF – inflected forms in the first position, if _BF – base forms preceded by inflected forms, if_IF – inflected forms preceded by inflected forms. TFT – total fixation time, GD – gaze duration, FFD – first fixation duration; SKIP – skipping rate; REG – regression rate. The upper-case represents either the unprimed conditions as in IF and BF, Lower case, if applicable, show the type of word form preceding the current word. stands for the prime.

		English L1				Dutch L1				English L2			
Variable	Type	N	Mean	SD	SE	N	Mean	SD	SE	N	Mean	SD	SE
TFT	BF	10143 (1122)	310.02	165.67	1.64	7456 (1001)	297.97	157.22	1.82	9855 (1063)	358.56	199.56	2.01
TFT	bf_BF	8734 (971)	301.40	158.51	1.70	6223 (847)	287.93	147.57	1.87	8502 (933)	347.88	191.31	2.07
TFT	if_BF	1263 (142)	284.84	147.57	4.15	1186 (189)	294.80	154.71	4.49	1288 (174)	323.19	173.52	4.83
TFT	IF	3978 (405)	329.47	171.78	2.72	3684 (497)	322.24	166.87	2.75	3815 (443)	387.56	206.27	3.34
TFT	if_IF	2715 (267)	311.72	161.36	3.10	2498 (329)	306.31	153.27	3.07	2527 (291)	367.17	189.71	3.77
TFT	bf_IF	1409 (155)	312.63	155.17	4.13	1233 (190)	306.88	159.58	4.54	1353 (182)	369.34	192.76	5.24
GD	BF	10143 (1122)	257.25	123.50	1.23	7456 (1001)	252.67	119.46	1.38	9855 (1063)	285.56	147.93	1.49
GD	bf_BF	8734 (971)	255.75	118.85	1.27	6223 (847)	250.31	115.67	1.47	8502 (933)	285.15	149.41	1.62
GD	if_BF	1263 (142)	243.22	108.91	3.06	1186 (189)	251.96	117.86	3.42	1288 (174)	261.87	126.26	3.52
GD	IF	3978 (405)	271.26	129.79	2.06	3684 (497)	264.61	126.26	2.08	3815 (443)	307.40	156.46	2.53
GD	if_IF	2715 (267)	264.34	121.34	2.33	2498 (329)	262.09	120.07	2.40	2527 (291)	302.66	152.62	3.04
GD	bf_IF	1409 (155)	265.66	115.62	3.08	1233 (190)	258.62	116.90	3.33	1353 (182)	306.01	149.64	4.07
FFD	BF	10143 (1122)	221.10	88.08	0.87	7456 (1001)	214.64	81.57	0.94	9855 (1063)	227.37	91.59	0.92
FFD	bf_BF	8734 (971)	220.45	84.91	0.91	6223 (847)	214.16	81.56	1.03	8502 (933)	230.34	94.13	1.02
FFD	if_BF	1263 (142)	219.76	86.34	2.43	1186 (189)	218.65	87.02	2.53	1288 (174)	227.92	89.90	2.50
FFD	IF	3978 (405)	226.83	89.88	1.42	3684 (497)	218.96	84.44	1.39	3815 (443)	235.49	96.28	1.56
FFD	if_IF	2715 (267)	223.51	82.15	1.58	2498 (329)	217.67	79.60	1.59	2527 (291)	235.73	94.20	1.87
FFD	bf_IF	1409 (155)	228.34	85.92	2.29	1233 (190)	219.67	79.13	2.25	1353 (182)	238.88	95.02	2.58
REG	BF	10143 (1122)	0.20	0.40	0.00	7456 (1001)	0.18	0.38	0.00	9855 (1063)	0.25	0.43	0.00
REG	bf_BF	8734 (971)	0.18	0.38	0.00	6223 (847)	0.15	0.36	0.00	8502 (933)	0.22	0.42	0.00
REG	if_BF	1263 (142)	0.17	0.38	0.01	1186 (189)	0.17	0.38	0.01	1288 (174)	0.22	0.41	0.01

REG	IF	3978 (405)	0.22	0.41	0.01	3684 (497)	0.21	0.41	0.01	3815 (443)	0.26	0.44	0.01
REG	if_IF	2715 (267)	0.18	0.39	0.01	2498 (329)	0.17	0.38	0.01	2527 (291)	0.22	0.42	0.01
REG	bf_IF	1409 (155)	0.18	0.38	0.01	1233 (190)	0.17	0.38	0.01	1353 (182)	0.22	0.42	0.01
SKIP	BF	15516 (1123)	0.31	0.46	0.00	11644 (1003)	0.28	0.45	0.00	13616 (1064)	0.27	0.44	0.00
SKIP	bf_BF	13439 (973)	0.32	0.47	0.00	9850 (849)	0.30	0.46	0.00	11799 (933)	0.27	0.45	0.00
SKIP	if_BF	1904 (142)	0.35	0.48	0.01	1802 (190)	0.31	0.46	0.01	1696 (174)	0.28	0.45	0.01
SKIP	IF	5581 (405)	0.26	0.44	0.01	5164 (498)	0.21	0.40	0.01	4833 (443)	0.21	0.41	0.01
SKIP	if_IF	3677 (267)	0.25	0.43	0.01	3362 (329)	0.23	0.42	0.01	3137 (291)	0.22	0.41	0.01
SKIP	bf_IF	2077 (155)	0.25	0.43	0.01	1794 (190)	0.24	0.43	0.01	1817 (183)	0.22	0.41	0.01

Note. The numbers in the parentheses indicate the number of unique words.

Table 2. Estimated identity (bf1-bf2 and if1-if2) and inflection (bf1-if2 and if1-bf2) priming effects. Highlighted are the contrasts that retain significance after the Bonferroni correction.

		English L1					Dutch L1					English L2				
DV	Predictor	Estimate	SE	T	p	p*	Estimate	SE	t	p	p*	Estimate	SE	t	p	p*
TFT	bf_BF	**-7.76**	**1.96**	**-3.96**	**0.000**	**0.000**	-4.75	2.14	-2.22	0.027	0.135	**-7.67**	**2.40**	**-3.20**	**0.001**	**0.005**
TFT	if_BF	-6.56	4.71	-1.39	0.164	0.820	2.95	4.21	0.70	0.484	1.000	**-15.33**	**5.44**	**-2.82**	**0.005**	**0.025**
TFT	if_IF	**-16.72**	**3.56**	**-4.70**	**0.000**	**0.000**	-7.41	3.54	-2.09	0.036	0.180	**-13.12**	**4.40**	**-2.98**	**0.003**	**0.015**
TFT	bf_IF	-10.91	5.35	-2.04	0.042	0.210	-11.03	5.16	-2.14	0.033	0.165	-7.68	6.13	-1.25	0.210	1.000
GD	bf_BF	-1.41	1.45	-0.98	0.329	1.000	-2.33	1.65	-1.42	0.156	0.780	-3.77	1.74	-2.17	0.030	0.150
GD	if_BF	-2.15	3.47	-0.62	0.536	1.000	-1.22	3.32	-0.37	0.713	1.000	**-9.91**	**3.84**	**-2.58**	**0.010**	**0.050**
GD	if_IF	-5.61	2.66	-2.11	0.035	0.175	-0.49	2.64	-0.19	0.853	1.000	-4.45	3.28	-1.36	0.175	0.875
GD	bf_IF	-0.93	3.91	-0.24	0.812	1.000	-5.63	3.82	-1.47	0.141	0.705	5.57	4.36	1.28	0.202	1.000
FFD	bf_BF	-0.14	1.10	-0.13	0.901	1.000	-0.86	1.23	-0.70	0.486	1.000	-0.18	1.22	-0.15	0.880	1.000
FFD	if_BF	-0.90	2.45	-0.37	0.713	1.000	-1.34	2.35	-0.57	0.569	1.000	0.04	2.61	0.02	0.987	1.000
FFD	if_IF	-1.71	1.92	-0.89	0.374	1.000	-1.48	1.88	-0.79	0.432	1.000	-0.16	2.24	-0.07	0.942	1.000
FFD	bf_IF	3.88	2.61	1.48	0.138	0.690	-0.10	2.59	-0.04	0.970	1.000	4.45	2.92	1.52	0.128	0.640
SKIP	bf_BF	0.04	0.03	1.25	0.212	1.000	0.03	0.04	0.93	0.350	1.000	0.07	0.03	2.05	0.040	0.200
SKIP	if_BF	-0.04	0.07	-0.62	0.539	1.000	0.14	0.07	2.08	0.038	0.190	-0.06	0.07	-0.86	0.388	1.000
SKIP	if_IF	-0.07	0.06	-1.17	0.244	1.000	**0.21**	**0.06**	**3.38**	**0.001**	**0.005**	0.06	0.06	1.01	0.310	1.000
SKIP	bf_IF	-0.09	0.08	-1.11	0.266	1.000	0.18	0.08	2.27	0.023	0.115	0.04	0.08	0.52	0.606	1.000
REG	bf_BF	**-0.12**	**0.04**	**-2.97**	**0.003**	**0.015**	-0.13	0.05	-2.54	0.011	0.055	-0.09	0.04	-2.25	0.025	0.125
REG	if_BF	-0.10	0.10	-1.06	0.287	1.000	0.08	0.10	0.84	0.399	1.000	-0.13	0.08	-1.53	0.127	0.635
REG	if_IF	**-0.23**	**0.07**	**-3.22**	**0.001**	**0.005**	**-0.22**	**0.07**	**-3.03**	**0.002**	**0.010**	**-0.19**	**0.07**	**-2.77**	**0.006**	**0.030**
REG	bf_IF	-0.24	0.10	-2.30	0.022	0.110	-0.23	0.10	-2.24	0.025	0.125	-0.16	0.09	-1.74	0.082	0.410

Note. (*) indicates Bonferroni corrected p value

Identity priming

We found considerable evidence in support of identity priming across samples and behavioral measures. Total fixation times to base forms following identical primes (bf – BF, tree-*tree*) decreased similarly in English L1 and L2, by 8 ms in each. The identity priming effect was also observed in the repetition of identical inflected forms (if – IF, trees-*trees*) in English L1 reading around 17 ms, and English L2 reading, around 13 ms. Surprisingly, identity priming was not observed in total fixation times in the Dutch L1 dataset.

Identity priming emerged in regression rates as well. The repetition of identical inflected forms (if – IF) resulted a similar effect in all, English L1, Dutch L1, and English L2, reading groups: Readers made 3% fewer regressions back to its second occurrence when they read the inflected form first (see Appendix, Table S5 to see estimated regression percentages for each condition). The English L1 group also benefited from reading a base identity prime (bf – BF): Regression rates back to a base form preceded by a base form decreased around 1%.

We also observed an identity priming effect in skip rates in Dutch L1 reading. When an inflected form occurred earlier in the text, skipping rate of its second occurrence increased by 4%.

Statistical power

In view of the (mostly null) results that our regression models revealed, it is important to evaluate the statistical power that the datasets at our disposal afford. We need to know the magnitude of the true population effect that the present data can detect with a reasonable likelihood, without committing the Type II error. It is worth noting that relatively few participants contributed to the GECO database: 14 L1 speakers of English and 19 L1 speakers of Dutch (which read half of the book in the original English and half in the Dutch translation).

Since we examine an existing corpus, our estimates of statistical power are based on observed data and fitted generalized linear mixed effects models (see Brysbaert, 2019; Brysbaert & Stevens, 2018 for worked examples). By varying the numeric contrast corresponding to the morphological priming effect, we determined the statistical power of detecting such a contrast with a nominal statistical significance (under 5%). We ran simulations that sampled observed data with replacement (100 iterations) and provided a statistical power estimate for the samples available for a specific dependent variable and priming type. For instance, the morphological priming effect of the if_BF type in English L1 total fixation time was estimated on the basis of 10,143 observations of unprimed base forms (BF) and 1,263 observations of base forms preceded by inflected forms, see Brysbaert and Stevens (2018) for detailed statistical procedure.

In the case of total fixation duration, simulations indicated across all samples that the available sample sizes are sufficient to detect a morphological priming effect of 20 ms with the acceptable power of roughly 80% (at the 5% level of significance). For instance, the estimated power of the bf_IF contrast was 82% [95% CI 69-91] in total fixation times in English L1 data. Using Westfall et al.'s (2014) technique for power analysis in mixed-effects models, we further determined that a 20 ms contrast corresponds to Cohen's *d* of 0.10, i.e., the effect size that is considered very small[2]. Similar procedures applied to other dependent variables determined very similar effect sizes (around *d* = 0.10) that can be detected with the acceptable statistical power in the available data. In other words, the book-length GECO corpus contains a sufficient amount of behavioral and linguistic material to avoid the Type II error even if the true population effect is very small.

What sample size would be necessary to provide sufficient statistical power for the observed morphological priming effects? Taking total fixation time as an example again, most numeric estimates in the observed data revolved around a 10 ms advantage in recognition speed to morphologically primed forms over unprimed ones (Table 2). This advantage corresponds to a negligibly small effect size estimate of Cohen's *d* = 0.05. Our power curve simulation indicated that a sample of 6,400 observations of morphologically primed forms is required to provide the 80% power (at the 5% level of significance) to detect an effect this small. This is four times the number of observations recommended by Brysbaert (2019) for a typical effect size in psychological research. Since the Agatha Christie's novel used in the GECO database contains roughly 150 morphologically primed lexical items of each type (see Table 1), to satisfy the power requirement above the novel is to be read by 43 readers. This number of readers is by no means excessive and the future development of long-text corpora might indeed reach the desirable statistical power to avoid the Type II error. Yet, with the estimated effect as weak as *d*=0.05, it is difficult to argue that such evidence would of any practical importance for accounts of morphological processing in natural reading.

General Discussion

The goal of this paper is to verify whether morphological priming – a reduction in the cognitive effort of processing a word if it is preceded by a morphologically related word – is observed in the natural reading of connected texts. Most experimental evidence on existence and nature of morphological priming has been obtained from the paradigms using isolated word recognition (e.g., lexical decision or fragment completion). Unsurprisingly, this evidence has been prominent in theories of morphological processing, see the Introduction. This paper tested whether morphological priming generalizes over paradigm- and task-specific demands of single word recognition. We examined long-term inflection priming patterns occurring when base forms and inflected forms of a word serve as the prime and the target in long natural texts. In the two languages under examination – English and Dutch – regular nominal and verbal inflection (tree – trees; or showed – showed) provides a maximum orthographic and semantic overlap with the base form. For this reason, of all types of

[2] If only within-section word repetitions were accepted, resulting sample sizes were smaller and provided the 80% power at the 5% level of significance to detect a 25 ms contrast, corresponding to *d* = 0.13.

morphological priming we expected inflection priming to elicit the strongest behavioral effects.

We made use of the GECO database (Cop et al., 2017) of the eye-tracking data recorded during book reading in English (by L1 and L2 readers) and Dutch (L2): this study has been modelled after a similar analysis of Spanish text reading by Kamienkowski et al. (2018). The GECO data contain hundreds of instances of both morphological priming (base forms preceding inflected forms or vice versa) and identity priming: virtually all are instances of long-term priming, with multiple intervening words. This makes the behavioral record of GEGO an excellent testbed for appraising whether morphological priming occurs under natural conditions of text reading for comprehension.

Morphological priming in L1

Main results of the present study can be summarized succinctly. We found no reliable statistical evidence for long-term inflection priming in English or Dutch L1 reading. These null results were true of the dependent variables representing all stages of the time course of word recognition, and for both types of inflection priming under consideration (base prime – inflected target and inflected prime – base target). At the same time, identity priming was robustly observed in multiple dependent variables across samples, suggesting that word repetition has an expected effect on reading speed (see the Introduction).

Further power analyses indicated that the observed failure to detect effects of morphological priming is unlikely to be a consequence of the Type II error. Available data allow for reliable detection of small effect sizes (around $d = 0.10$ for all dependent variables). The observed non-significant priming effects were negligibly small (around $d = 0.05$). A four-fold increase in the amount of data would be required to provide sufficient power to render such effects significant (with the nominal likelihood and significance threshold). While such an effort is possible, it will not change the observation that the practical importance of morphological priming in natural reading of long texts for comprehension is minute. This analysis converges with an earlier demonstration by Schmidtke et al. (2020) that very few of the morpho-semantic effects that lexical decision studies identified as relevant for visual recognition of compounds are detectable in a large-scale corpus of eye-movements recorded during sentence reading.

Amenta and Crepaldi's (2012) review states: "Morphological priming has been so extensively observed […] that it does not make any sense to ask ourselves whether it exists or not". As the review indicates, this appraisal arises from a survey of work done in single word recognition paradigms. Our data cast doubt on generalizability of this statement over English and Dutch natural text reading in L1, while null results in Kamienkowski et al. (2018) make the statement problematic for Spanish also. It is worth noting that, collectively, data from English, Dutch and Spanish L1 readers likely account for the lion's share of empirical evidence upon which current theoretical accounts of morphological processing are built. The null findings in this study further highlight the need for psycholinguistic research to examine natural reading behaviour. The present results – especially if confirmed in future work – indicate the need to revise those accounts to align with behaviors shown naturally, outside of demands and artificial constraints of laboratory tasks.

Morphological priming in L2

The only dataset in which one type of morphological priming showed statistical significance was that with Dutch students reading in English (their L2). Base forms primed by inflected forms (trees – tree) showed a reliable advantage over unprimed base forms in gaze duration (10 ms) and total fixation duration (15 ms). In the Introduction, we offered a brief overview of the L2 morphological priming literature using isolated word presentation (based on Jacob, 2018). Part of the literature suggests that inflection priming is not found among L2 readers (e.g., Clahsen & Neubauer, 2010; Jacob, Fleischhauer, & Clahsen, 2013; Jacob et al., 2017; Kirkici & Clahsen, 2013; Neubauer & Clahsen, 2009; Silva & Clahsen, 2008), while other sources find that L2 and L1 readers demonstrate equally strong inflection priming (Basnight-Brown et al., 2007; Coughlin & Tremblay, 2014; De Grauwe et al., 2014; Diependaele et al., 2011; Feldman et al., 2010; Foote, 2015). Our data point to a third scenario: L2 readers of English show long-term inflection priming when L1 readers do not. One possible explanation of this discrepancy is of a statistical nature: since L2 readers have longer reading times than L1 readers, all effects on their reading times are also larger and easier to reach inferential significance in statistical models (see Jacob, 2018). We are not aware of any theoretical accounts that predict stronger morphological priming effects among L2 readers compared to L1 readers, or a reason for why the base form primed by an inflected form shows the speed advantage, but an inflected form primed by a base form does not, when compared to the respective unprimed forms.

Identity priming in L1 and L2

Our findings showed that identity priming could emerge in long text reading even when the repeated words are far apart from each other (e.g., the median lag in the English texts was 6,357 and 93% of lags were over 100 words, and the lag distribution in Dutch was similar). This effect was observed both in L1 and L2 English reading. Nearly all supporting evidence came from late eye-tracking measures (regression rates and total fixation duration). These results differ from the findings of Kamienkowski et al. (2018), who reported the influence of repeated words on early reading measures like first fixation duration and gaze duration. One difference between these two studies is the structure of the selected texts: Kamienkowski et al. (2018) analyzed multiple texts that were relatively short, whereas we used long, single connected texts. Given that both studies did not find any effect of distance on reading time, the length of the text or the time elapsed between the repeated words is not likely to cause this difference.

Another difference between the two studies are the languages examined having varying degrees of orthographic transparency. An argument can be made that in highly transparent languages (Spanish), the likelihood of finding the effects in early measures increases, and in less transparent languages, the emergence of the effects is likely to shift to late measures, as in English, and Dutch falling somewhere in the middle, showing evidence from both measures.

Another possibility could be attributed to the predictability of a word in a given context. Readers tend to generate predictions about upcoming words. Several studies have demonstrated that readers spend less time on words that are highly predictable (see Lowder, Choi, Ferreira, & Henderson, 2018 for a review). Similarly, Kamienkowski et al. (2018) associated faster reading time in the first pass reading with higher lexical predictability. It is important to note, however, that as the genre of the analyzed text is detective fiction in our study, unexpected twists and turns are parts of the nature of this genre which might reduce predictability in general. Altogether, our findings suggest that word reading is influenced by different levels of information. Prior exposure to a word seems to consistently reduce the overall reading time for subsequent occurrences. Yet, the contextual level of information seems to play a global role in the time course of the word reading process. Future studies should elaborate on these findings by looking at the word- and context-level factors together.

Limitations and Future Directions.

Ecological validity of natural reading afforded by eye-tracking studies of book reading for comprehension enables researchers to avoid many pitfalls of other paradigms (see the Introduction) but is not without a cost. Natural texts are not experimentally controlled for the number of cases of priming, their distribution over text, lexical characteristics of primes and targets (e.g., word length and frequency), the number and nature of words intervening between the prime and the target and other factors of potential relevance. Tightly controlled experiments create highly artificial reading conditions but eliminate these sources of variance. Thus, results of any corpus study, including the present one, are correlational. While great care was taken to reduce undesirable variance through exclusion of outlier stimuli and statistical control of covariates, a possibility remains that a specific effect (including a null effect) is partly due to unaccounted extraneous factors, unrelated to the factor of interest. What reduces this possibility is replication of the central result in multiple corpora: there is no reason to assume that unrelated corpora would give rise to the same values and distributions of extraneous factors leading to the same bias in behavioral outcomes. The present English and Dutch L1 data replicate the null effects of long-term inflection priming reported for Spanish (Kamienkowski et al., 2018). We encourage additional replication in these and other languages, which would make the present set of findings more robust. It is also worth mentioning that the current study examined the priming effect on the first and second occurrences of a base and inflected form. It is possible that morphological priming in L1 reading starts to emerge after the second occurrence. To address this issue, we also encourage future research to incorporate subsequent instances of the base and target words.

Current results indicate morphological priming in L2 as a promising direction for future research. To our knowledge, few studies have been dedicated to L1-L2 comparisons in long-term inflection priming. This study justifies a greater theoretical attention to such a comparison.

We chose inflection priming in regular nominal and verbal forms because it was most likely to elicit effects, due to a substantial orthographic and semantic overlap between base and inflected forms. It is worthwhile to conduct similar studies on irregular inflection, derivation, and compounding, including consideration of semantic variables in the latter two types of morphological productivity. We acknowledge that the purely orthographically related prime-target pairs (e.g., *scandal – scan*) were not addressed in the current study. As literature has suggested, L2 speakers might also be likely to rely more on surface form characteristics of a word than L1 speakers in long text reading. Future studies should also include an orthographic control set to be able to determine to what extent priming effects result from purely orthographic overlap in natural reading of long texts. Corpus studies like the present one offer an easy and informative way for such sweeping examination of morphological phenomena in a few most studied languages.

References

Amenta, S. & Crepaldi, D. (2012). Morphological Processing as We Know It: An Analytical Review of Morphological Effects in Visual Word Identification. *Frontiers in Psychology*, 3.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8. Retrieved from http://CRAN.R-project.org/package=lme4

Bertram, R. (2011). Eye movements and morphological processing in reading. *The Mental Lexicon*, *6*(1), 83-109.

Basnight-Brown, D., Chen, L., Hua, S., Kostić, A. and Feldman, L. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of Memory and Language*, 57(1), pp.65-80.

Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *The Quarterly Journal of Experimental Psychology*, *42*(4), 693-711.

Boston, M. F., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, *2*(1). https://doi.org/10.16910/jemr.2.1.1

Burani, C. (2010). Word morphology enhances reading fluency in children with developmental dyslexia. *Lingue e linguaggio*, *9*(2), 177-198.

Burani, C., Marcolini, S., De Luca, M., & Zoccolotti, P. (2008). Morpheme-based reading aloud: Evidence from dyslexic and skilled Italian readers. *Cognition*, *108*(1), 243-262.

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, *44*(4), 991–997. https://doi.org/10.3758/s13428-012-0190-4

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 16. DOI: http://doi.org/10.5334/joc.72

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(1), 9. https://doi.org/10.5334/joc.10

Carlisle, J. F., & Stone, C. A. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly, 40*(4), 428-449.

Chamberland, C., Saint-Aubin, J., & Légère, M. A. (2013). The impact of text repetition on content and function words during reading: Further evidence from eye movements. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *67*(2), 94.

Ciaccio, L. A., & Jacob, G. (2019). Native speakers like affixes, L2 speakers like letters? An overt visual priming study investigating the role of orthography in L2 morphological processing. *PLOS ONE*, *14*(12), e0226482. https://doi.org/10.1371/journal.pone.0226482

Clahsen, H. and Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua*, 120(11), pp.2627-2637. http://dx.doi.org/10.1017/S1366728914000662

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

Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. Second Language Research, 32(3), 453–467. https://doi.org/10.1177/0267658316637401

Cop, U., Dirix, N., Drieghe, D., & Duyck, W. (2017). Presenting GECO: An eyetracking corpus of monolingual and bilingual sentence reading. *Behavior Research Methods*, *49*(2), 602-615.

Coughlin, C. And Tremblay, A. (2014). Morphological decomposition in native and non-native French speakers. *Bilingualism: Language and Cognition*, 18(3), pp.524-542.

Crepaldi, D., Rastle, K., Coltheart, M. and Nickels, L. (2010). 'Fell' primes 'fall', but does 'bell' prime 'ball'? Masked priming with irregularly-inflected primes. *Journal of Memory and Language*, 63(1), 83-99.

Dann, K., Veldre, A., & Andrews, S. (2018). *Morphological preview effects in English are restricted to suffixed words*. *Ea 3742*, 1–57. https://doi.org/10.1037/xlm0001029

De Grauwe, S., Lemhöfer, K., Willems, R. and Schriefers, H. (2014). L2 speakers decompose morphologically complex verbs: fMRI evidence from priming of transparent derived verbs. *Frontiers in Human Neuroscience*, 8.

Deacon, S. H., Tong, X., & Mimeau, C. (2019). 15 Morphological and Semantic Processing in Developmental Dyslexia. *Developmental dyslexia across languages and writing systems*, *327*.

Diependaele, K., Duñabeitia, J., Morris, J. and Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, 64(4), pp.344-358.

Dirix, N., Brysbaert, M., & Duyck, W. (2019). How well do word recognition measures correlate? Effects of language context and repeated presentations. *Behavior Research Methods*, *51*(6), 2800-2816.

Feldman, L. B., & Andjelković, D. (1992). Morphological analysis in word recognition. In *Advances in psychology* (Vol. 94, pp. 343-360). North-Holland.

Feldman, L. B. (1992). Bi-alphabetism and the design of a reading mechanism. In D. M. Willows, R. S. Kruk, & E. Corcos (Eds.), Visual processes in reading and reading disabilities. Hillsdale, NJ: Erlbaum.

Feldman, L. B., & Bentin, S. (1994). Morphological analysis of disrupted morphemes: Evidence from Hebrew. *The Quarterly Journal of Experimental Psychology*, *47*(2), 407-435.

Feldman, L. B., Kostić, A., Basnight-Brown, D. M., Durđević, D. F., & Pastizzo, M. J. (2010). Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism (Cambridge, England)*, *13*, 119–135. https://doi.org/10.1017/S1366728909990459

Feldman, L. B., Rueckl, J., Diliberto, K., Pastizzo, M., & Vellutino, F. R. (2002). Morphological analysis by child readers as revealed by the fragment completion task. *Psychonomic Bulletin & Review*, *9(3)*, 529-535.

Foote, R. (2015). The Storage And Processing Of Morphologically Complex Words In L2 Spanish. *Studies in Second Language Acquisition*, 39(4), 735-767.

Forster, K., Davis, C., Schoknecht, C. and Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation?. *The Quarterly Journal of Experimental Psychology Section A*, 39(2), 211-251.

Frost, R., Forster, K. and Deutsch, A. (1997). "What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation": Correction to Frost et al. 1997. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(5), 1189-1191.

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., Yoon, H. (2017). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition, 21*(3), 1–22. https://doi:10.1017/S1366728917000219

Gonnerman, L., Seidenberg, M. and Andersen, E., 2007. Graded semantic and phonological similarity effects in priming: Evidence for a distributed connectionist approach to morphology. *Journal of Experimental Psychology: General*, 136(2), pp.323-345.

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., & Fagot, J. (2012). Orthographic processing in baboons (Papio papio). *Science*, *336*(6078), 245-248.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychological Review*, *103*(3), 518.

Grainger, J., Colé, P. and Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, 30(3), 370-384.

Green, P. and MacLeod, C.J. (2016), SIMR: an R package for power analysis of generalized linear mixed models by simulation. Methods Ecol Evol, 7: 493-498. https://doi.org/10.1111/2041-210X.12504

Heyer, V. And Clahsen, H. (2015). Late bilinguals see a scan in scanner and in scandal: dissecting formal overlap from morphological priming in the processing of derived words. *Bilingualism: Language and Cognition*, 18(3), 543-550.

Jacob, G., Fleischhauer, E. And Clahsen, H. (2013). Allomorphy and affixation in morphological processing: A cross-modal priming study with late bilinguals. *Bilingualism: Language and Cognition*, 16(4), 924-933.

Jacob, G., Heyer, V. and Veríssimo, J. (2017). Aiming at the same target: A masked priming study directly comparing derivation and inflection in the second language. *International Journal of Bilingualism*, 22(6), 619-637.

Jacob, G. (2018). Morphological priming in bilingualism research. *Bilingualism: Language and Cognition*, *21*(3), 443-447.

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology*, *94*(2), 223-244.

Kamienkowski, J. E., Carbajal, M. J., Bianchi, B., Sigman, M., & Shalom, D. E. (2018). Cumulative repetition effects across multiple readings of a word: Evidence from eye movements. *Discourse Processes*, *55*(3), 256-271.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650. https://doi.org/10.3758/BRM.42.3.643

Kirkici, B., & Clahsen, H. (2013). Inflection and derivation in native and non-native language processing: Masked priming experiments on Turkish. *Bilingualism*, *16*(4), 776.

Kuperman, V., Matsuki, K., & Van Dyke, J. A. (2018). Contributions of reader- and text-level characteristics to eye-movement patterns during passage reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(11), 1687.

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *Quarterly Journal of Experimental Psychology*, *66*(3), 563-580.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: a quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Leminen, A., Smolka, E., Duñabeitia, J. and Pliatsikas, C. (2019). Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding). *Cortex*, 116, 4-44.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 381–390. https://doi.org/10.1037/a0014875

Mann, V., & Singson, M. (2003). Linking morphological knowledge to English decoding ability: Large effects of little suffixes. In *Reading complex words* (pp. 1-25). Springer, Boston, MA.

Marelli, M., Traficante, D., & Burani, C. (2020). Reading morphologically complex words: Experimental evidence and learning models. *Word Knowledge and Word Usage*, 553.

Marelli, M., & Luzzatti, C. (2012). Frequency effects in the processing of Italian nominal compounds: Modulation of headedness and semantic transparency. *Journal of Memory and Language*, *66*(4), 644-664.

Marslen-Wilson, W., Tyler, L., Waksler, R. and Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101(1), 3-33.

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

Mousikou, P., & Schroeder, S. (2019). Morphological processing in single-word and sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(5), 881.

Münte, T. F., Say, T., Clahsen, H., Schiltz, K., & Kutas, M. (1999). Decomposition of morphologically complex words in English: Evidence from event-related brain potentials. *Cognitive Brain Research*, *7*(3), 241-253.

Linke, M., Bröker, F., Ramscar, M., & Baayen, H. (2017). Are baboons learning “orthographic” representations? Probably not. *PLOS ONE*, *12*(8), e0183876. https://doi.org/10.1371/journal.pone.0183876

Lensink, S. E., Verdonschot, R. G., & Schiller, N. O. (2014). Morphological priming during language switching: an ERP study. *Frontiers in Human Neuroscience*, *8*, 995.

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. R package version,

Version: 1.5.2 Available at: https://CRAN.R-project.org/package=emmeans.

Linke, M., Bröker, F., Ramscar, M., & Baayen, H. (2017). Are baboons learning "orthographic" representations? Probably not. *PloS one*, *12*(8), e0183876.

Liversedge, S. P., Blythe, H. I., & Drieghe, D. (2012). Beyond isolated word recognition. *Behavioral and Brain Sciences*, *35*(5), 293-294.

Lowder, M. W., Choi, W., Ferreira, F., & Henderson, J. M. (2018). Lexical predictability during natural reading: Effects of surprisal and entropy reduction. *Cognitive Science, 42(Suppl 4), 1166–1183. https://doi.org/10.1111/cogs.12597*

Marslen-Wilson, W. D. (2007). Processes in language comprehension. *Oxford handbook of psycholinguistics*, 495-524.

Masson, M. E. J., & MacLeod, C. M. (1992). Reenacting the route to interpretation: Enhanced perceptual identification without prior perception. *Journal of Experimental Psychology: General, 121(2)*, 145–176. https://doi.org/10.1037/0096-3445.121.2.145

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychological Bulletin*, *81*(12), 899.

Milin, P., Smolka, E., & Feldman, L. B. (2017). Models of lexical access and morphological processing. In E. M. Fernández & H. S. Cairns (Eds.), *The handbook of psycholinguistics* (pp. 240–268). Malden: Wiley-Blackwell.

Mousikou, P., & Schroeder, S. (2019). Morphological processing in single-word and sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(5), 881.

Neubauer, K. and Clahsen, H. (2009). Decomposition of inflected words in a second language. *Studies in Second Language Acquisition*, 31(3), pp.403-435.

Paterson, K. B., Alcock, A., & Liversedge, S. P. (2011). Morphological priming during reading: Evidence from eye movements. *Language and Cognitive Processes*, *26*(4-6), 600-623.

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. *Psychological Research*, *72*(6), 675-688.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *49*(2), 151.

Rastle, K., Davis, M., Marslen-Wilson, W. and Tyler, L. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, 15(4-5), 507-537.

Raveh, M. and Rueckl, J. (2000). Equivalent Effects of Inflected and Derived Primes: Long-Term Morphological Priming in Fragment Completion and Lexical Decision. *Journal of Memory and Language*, 42(1), 103-119.

Rayner, K., & Liversedge, S. P. (2011). *Linguistic and cognitive influences on eye movements during reading.* In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *Oxford library of psychology. The Oxford handbook of eye movements* (p. 751–766). Oxford University Press.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K. (2009) Eye movements and attention in reading, scene perception, and visual search, *The Quarterly Journal of Experimental Psychology*, 62:8, 1457-1506, DOI: 10.1080/17470210902816461

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.

Rayner, K., Raney, G. E., & Pollatsek, A. (1995). Eye movements and discourse processing. In R. Lorch & E. O'Brien (Eds.), *Sources of coherence in reading* (pp. 9–36). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65-81.

Rodriguez-Fornells, A., Münte, T. and Clahsen, H. (2002). Morphological Priming in Spanish Verb Forms: An ERP Repetition Priming Study. *Journal of Cognitive Neuroscience*, 14(3), 443-454.

Royle, P., Drury, J. E., Bourguignon, N., & Steinhauer, K. (2012). The temporal dynamics of inflected word recognition: A masked ERP priming study of French verbs. *Neuropsychologia*, *50*(14), 3542-3553.

Rueckl, J. & Aicher, K. (2008). Are CORNER and BROTHER morphologically complex? Not in the long term. *Language and Cognitive Processes*, 23(7-8), 972-1001.

Rueckl, J. G., & Galantucci, B. (2005). The locus and time course of long-term morphological priming. *Language and Cognitive Processes, 20(1-2)*, 115-138.

Rueckl, J. G., Mikolinski, M., Raveh, M., Miner, C. S., & Mars, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language, 36(3)*, 382-405.

Schmidtke, D., Van Dyke, J. A., & Kuperman, V. (2020). CompLex: an eye-movement database of compound word reading in English. *Behavior Research Methods*, 53(1), 59–77. https://doi.org/10.3758/s13428-020-01397-1

Schmidtke, D., & Kuperman, V. (2020). Psycholinguistic Methods and Tasks in Morphology. In *Oxford Research Encyclopedia of Linguistics*.

Schmidtke, D., & Kuperman, V. (2019). A paradox of apparent brainless behavior: The time-course of compound word recognition. *Cortex*, *116*, 250-267.

Schmidtke, D., Matsuki, K., & Kuperman, V. (2017). Surviving blind decomposition: A distributional analysis of the time-course of complex word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1793.

Silva, R. And Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11(2), 245-260.

Stanners, R., Neiser, J., Hernon, W. and Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, 18(4), 399-412.

Traficante, A., Calzoletti, L., Veneziani, M., Ali, B., De Gasperis, G., Di Giorgio, A. M., ... & Schisano, E. (2011). Data reduction pipeline for the Hi-GAL survey. *Monthly Notices of the Royal Astronomical Society*, *416*(4), 2932-2943.

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045. https://doi.org/10.1037/xge0000014

Weyerts, H., Münte, T. F., Smid, H. G., & Heinze, H. J. (1996). Mental representations of morphologically complex words: an event-related potential study with adult humans. *Neuroscience Letters*, *206*(2-3), 125-128.

Von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, *94*, 119-13

Supplementary Materials

Table S1. Results of the linear mixed-effects models fitted to Total Fixation Times for English L1, Dutch L1, and English L2. Parentheses indicate reference levels.

	English L1 (BF)			English L1 (IF)			Dutch L1 (BF)			Dutch L1 (IF)			English L2 (BF)			English L2 (IF)		
Fixed Effects	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*
Intercept	268.34	13.95	**<0.001**	248.79	21.01	**<0.001**	266.28	12.88	**<0.001**	256.60	15.78	**<0.001**	339.34	16.31	**<0.001**	312.30	26.12	**<0.001**
bf_BF	-7.76	1.96	**<0.001**				-4.75	2.14	**0.027**				-7.67	2.40	**0.001**			
if_BF	-6.56	4.71	0.164				2.95	4.21	0.484				-15.33	5.44	**0.005**			
Ordinal Position	-0.96	1.25	0.440	-1.97	2.12	0.353	-7.22	2.49	**0.004**	-10.52	3.96	**0.008**	-11.07	3.04	**<0.001**	-14.93	4.94	**0.003**
Noun	-3.43	2.82	0.223	-6.81	14.76	0.644	-4.11	3.14	0.190	-9.18	5.92	0.121	-12.17	3.74	**0.001**	-16.25	17.57	0.355
Verb	-3.82	4.23	0.367	1.97	14.51	0.892	-7.02	4.13	0.089	-0.73	4.93	0.882	6.08	5.46	0.265	0.89	17.28	0.959
Length	8.49	0.63	**<0.001**	11.12	1.15	**<0.001**	7.46	0.55	**<0.001**	8.66	0.88	**<0.001**	12.44	0.85	**<0.001**	13.95	1.46	**<0.001**
Frequency (log)	-5.46	0.74	**<0.001**	-4.72	1.11	**<0.001**	-7.05	0.76	**<0.001**	-5.50	1.02	**<0.001**	-11.54	0.95	**<0.001**	-9.35	1.41	**<0.001**
bf_IF				-10.91	5.35	**0.041**				-11.03	5.16	**0.033**				-7.68	6.13	0.210
if_IF				-16.72	3.56	**<0.001**				-7.41	3.54	**0.036**				-13.12	4.40	**0.003**
Random Effects																		
σ^2	14017.23			15339.60			12527.28			14843.23			21116.70			23069.91		
τ_{00}	888.75 $_{WORD}$			1030.19 $_{WORD}$			698.96 $_{WORD}$			1190.63 $_{WORD}$			1484.59 $_{WORD}$			1647.04 $_{WORD}$		
	1751.82 $_{PP\_NR}$			1801.45 $_{PP\_NR}$			1710.23 $_{PP\_NR}$			2060.33 $_{PP\_NR}$			2651.47 $_{PP\_NR}$			3723.11 $_{PP\_NR}$		
N	1264 $_{WORD}$			559 $_{WORD}$			1120 $_{WORD}$			655 $_{WORD}$			1171 $_{WORD}$			585 $_{WORD}$		
	14 $_{PP\_NR}$			14 $_{PP\_NR}$			18 $_{PP\_NR}$			18 $_{PP\_NR}$			19 $_{PP\_NR}$			19 $_{PP\_NR}$		
Observations	19493			7850			14378			7186			19035			7478		
Marginal R^2 / Conditional R^2	0.041 / 0.193			0.036 / 0.186			0.056 / 0.208			0.047 / 0.218			0.075 / 0.227			0.050 / 0.229		

Table S2. Results of the linear mixed-effects models fitted to Gaze Durations for English L1, Dutch L1, and English L2. Parentheses indicate reference levels.

	English L1 (BF)			English L1 (IF)			Dutch L1 (BF)			Dutch L1 (IF)			English L2 (BF)			English L2 (IF)		
Fixed Effects	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*
Intercept	228.85	11.32	**<0.001**	222.68	16.12	**<0.001**	238.92	10.52	**<0.001**	226.37	12.35	**<0.001**	253.12	13.05	**<0.001**	270.14	19.66	**<0.001**
bf_BF	-1.41	1.45	0.329				-2.33	1.65	0.156				-3.77	1.74	**0.030**			
if_BF	-2.15	3.47	0.536				-1.22	3.32	0.713				-9.91	3.84	**0.010**			
Ordinal Position	0.24	0.92	0.797	-1.30	1.55	0.400	0.94	1.97	0.632	-2.46	2.96	0.407	1.51	2.17	0.486	-7.23	3.56	**0.042**
Noun	0.27	2.07	0.895	-6.56	10.82	0.544	-5.08	2.53	**0.045**	-5.02	4.37	0.251	-5.89	2.54	**0.020**	-29.54	12.22	**0.016**
Verb	-0.52	3.11	0.867	-5.80	10.63	0.585	-4.84	3.32	0.145	-0.36	3.64	0.921	-5.57	3.73	0.136	-16.86	12.02	0.161
Length	5.44	0.46	**<0.001**	7.48	0.84	**<0.001**	4.89	0.44	**<0.001**	5.34	0.65	**<0.001**	8.82	0.58	**<0.001**	10.06	0.99	**<0.001**
Frequency (log)	-3.45	0.54	**<0.001**	-3.25	0.81	**<0.001**	-5.18	0.61	**<0.001**	-3.75	0.75	**<0.001**	-5.17	0.65	**<0.001**	-6.07	0.96	**<0.001**
bf_IF				-0.93	3.91	0.812				-5.63	3.82	0.141				5.57	4.36	0.201
if_IF				-5.61	2.65	**0.035**				-0.49	2.64	0.853				-4.45	3.28	0.175
Random Effects																		
σ^2	7735.95			8685.28			7346.02			8310.87			11308.90			13188.26		
τ_{00}	472.67 $_{\text{WORD}}$			507.41 $_{\text{WORD}}$			527.72 $_{\text{WORD}}$			632.01 $_{\text{WORD}}$			596.68 $_{\text{WORD}}$			583.79 $_{\text{WORD}}$		
	1271.04 $_{\text{PP\_NR}}$			1299.51 $_{\text{PP\_NR}}$			1157.26 $_{\text{PP\_NR}}$			1428.69 $_{\text{PP\_NR}}$			2119.84 $_{\text{PP\_NR}}$			2942.56 $_{\text{PP\_NR}}$		
N	1264 $_{\text{WORD}}$			559 $_{\text{WORD}}$			1120 $_{\text{WORD}}$			655 $_{\text{WORD}}$			1171 $_{\text{WORD}}$			585 $_{\text{WORD}}$		
	14 $_{\text{PP\_NR}}$			14 $_{\text{PP\_NR}}$			18 $_{\text{PP\_NR}}$			18 $_{\text{PP\_NR}}$			19 $_{\text{PP\_NR}}$			19 $_{\text{PP\_NR}}$		
Observations	19647			7896			14499			7229			19135			7507		
Marginal R^2 / Conditional R^2	0.029 / 0.208			0.025 / 0.193			0.045 / 0.223			0.031 / 0.223			0.051 / 0.235			0.038 / 0.241		

Table S3. Results of the generalized linear mixed-effects models fitted to Regressions for English L1, Dutch L1, and English L2. Parentheses indicate reference levels.

	English L1 (BF)			English L1 (IF)			Dutch L1 (BF)			Dutch L1 (IF)			English L2 (BF)			English L2 (IF)		
Fixed Effects	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*
Intercept	-1.81	0.22	**<0.001**	-2.12	0.38	**<0.001**	-1.88	0.21	**<0.001**	-1.69	0.23	**<0.001**	-1.06	0.19	**<0.001**	-1.39	0.33	**<0.001**
bf_BF	-0.12	0.04	**0.003**				-0.13	0.05	**0.011**				-0.09	0.04	**0.025**			
if_BF	-0.10	0.10	0.287				0.08	0.10	0.399				-0.13	0.08	0.127			
Ordinal Position	-0.08	0.02	**0.002**	-0.02	0.04	0.571	-0.18	0.05	**<0.001**	-0.15	0.06	**0.017**	-0.21	0.04	**<0.001**	-0.10	0.07	0.118
Noun	-0.04	0.05	0.508	0.16	0.30	0.588	0.03	0.07	0.716	-0.19	0.11	0.075	-0.13	0.05	**0.009**	-0.04	0.25	0.883
Verb	-0.06	0.08	0.478	0.36	0.30	0.223	-0.15	0.09	0.104	-0.14	0.09	0.109	0.18	0.08	**0.020**	0.06	0.24	0.805
Length	0.06	0.01	**<0.001**	0.06	0.02	**0.006**	0.05	0.01	**<0.001**	0.05	0.02	**0.001**	0.04	0.01	**<0.001**	0.04	0.02	**0.030**
Frequency (log)	-0.03	0.01	0.057	-0.02	0.02	0.381	-0.04	0.02	**0.017**	-0.02	0.02	0.247	-0.06	0.01	**<0.001**	-0.03	0.02	0.117
bf_IF				-0.24	0.10	**0.022**				-0.23	0.10	**0.025**				-0.15	0.09	0.082
if_IF				-0.23	0.07	**0.001**				-0.22	0.07	**0.002**				-0.19	0.07	**0.006**
Random Effects																		
σ^2	3.29			3.29			3.29			3.29			3.29			3.29		
τ_{00}	0.24 $_{\text{WORD}}$			0.26 $_{\text{WORD}}$			0.24 $_{\text{WORD}}$			0.26 $_{\text{WORD}}$			0.18 $_{\text{WORD}}$			0.18 $_{\text{WORD}}$		
	0.30 $_{\text{PP\_NR}}$			0.34 $_{\text{PP\_NR}}$			0.18 $_{\text{PP\_NR}}$			0.11 $_{\text{PP\_NR}}$			0.20 $_{\text{PP\_NR}}$			0.29 $_{\text{PP\_NR}}$		
N	1264 $_{\text{WORD}}$			559 $_{\text{WORD}}$			1120 $_{\text{WORD}}$			655 $_{\text{WORD}}$			1171 $_{\text{WORD}}$			585 $_{\text{WORD}}$		
	14 $_{\text{PP\_NR}}$			14 $_{\text{PP\_NR}}$			18 $_{\text{PP\_NR}}$			18 $_{\text{PP\_NR}}$			19 $_{\text{PP\_NR}}$			19 $_{\text{PP\_NR}}$		
Observations	20140			8102			14865			7415			19645			7695		
Marginal R^2 / Conditional R^2	0.010 / 0.148			0.010 / 0.163			0.019 / 0.129			0.017 / 0.116			0.022 / 0.123			0.009 / 0.133		

Table S4. Results of the generalized linear mixed-effects models fitted to Skips for English L1, Dutch L1, and English L2. Parentheses indicate reference levels.

	English L1 (BF)			English L1 (IF)			Dutch L1 (BF)			Dutch L1 (IF)			English L2 (BF)			English L2 (IF)		
Fixed Effects	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*
Intercept	0.35	0.16	**0.024**	-0.56	0.30	0.063	0.46	0.17	**0.005**	0.29	0.22	0.203	0.20	0.19	0.283	-0.27	0.30	0.369
bf_BF	0.04	0.03	0.212				0.03	0.03	0.350				0.07	0.03	**0.040**			
if_BF	-0.04	0.07	0.539				0.14	0.07	**0.038**				-0.06	0.07	0.388			
Ordinal Position	0.02	0.02	0.383	0.02	0.03	0.503	0.11	0.04	**0.003**	0.02	0.06	0.803	-0.05	0.04	0.240	-0.02	0.06	0.790
Noun	-0.03	0.04	0.466	0.96	0.24	**<0.001**	0.02	0.05	0.736	0.16	0.09	0.068	0.04	0.04	0.376	0.21	0.20	0.300
Verb	0.01	0.06	0.867	0.78	0.24	**0.001**	0.01	0.06	0.854	-0.06	0.08	0.415	-0.06	0.06	0.358	0.11	0.20	0.573
Length	-0.23	0.01	**<0.001**	-0.19	0.02	**<0.001**	-0.26	0.01	**<0.001**	-0.25	0.02	**<0.001**	-0.26	0.01	**<0.001**	-0.19	0.02	**<0.001**
Frequency (log)	0.04	0.01	**<0.001**	0.01	0.02	0.513	0.04	0.01	**<0.001**	0.03	0.02	0.073	0.03	0.01	**0.015**	0.01	0.02	0.570
bf_IF				-0.09	0.08	0.266				0.18	0.08	**0.023**				0.04	0.08	0.606
if_IF				-0.07	0.06	0.244				0.21	0.06	**0.001**				0.06	0.06	0.310
Random Effects																		
σ^2	3.29			3.29			3.29			3.29			3.29			3.29		
τ_{00}	0.14 $_{WORD}$			0.22 $_{WORD}$			0.10 $_{WORD}$			0.19 $_{WORD}$			0.11 $_{WORD}$			0.13 $_{WORD}$		
	0.15 $_{PP\_NR}$			0.16 $_{PP\_NR}$			0.20 $_{PP\_NR}$			0.29 $_{PP\_NR}$			0.36 $_{PP\_NR}$			0.46 $_{PP\_NR}$		
N	1265 $_{WORD}$			559 $_{WORD}$			1123 $_{WORD}$			656 $_{WORD}$			1172 $_{WORD}$			585 $_{WORD}$		
	14 $_{PP\_NR}$			14 $_{PP\_NR}$			18 $_{PP\_NR}$			18 $_{PP\_NR}$			19 $_{PP\_NR}$			19 $_{PP\_NR}$		
Observations	30859			11335			23296			10320			27111			9787		
Marginal R^2 / Conditional R^2	0.082 / 0.156			0.040 / 0.139			0.133 / 0.207			0.103 / 0.217			0.088 / 0.202			0.035 / 0.181		

Table S5. Conditional means estimated based on regression models for the six levels of factor Type, i.e., the six categories formed by the type of the repeated word (base form bf of inflected form if), and its position in the pair of occurrences (first or second). CI stands for the 95% confidence interval. TFT – total fixation time, GD – gaze duration, FFD – first fixation duration; SKIP – skipping rate; REG – regression rate.

		English L1			Dutch L1			English L2		
Variable	Type	Mean	Lower CI	Upper CI	Mean	Lower CI	Upper CI	Mean	Lower CI	Upper CI
TFT	BF	289.51	267.32	311.71	274.62	255.18	294.06	338.30	314.74	361.86
TFT	bf_BF	281.76	259.51	304.00	269.87	250.37	289.37	330.63	307.00	354.27
TFT	if_BF	282.95	259.41	306.49	277.57	257.04	298.10	322.98	297.75	348.20
TFT	IF	308.89	284.24	333.53	298.02	276.36	319.68	360.79	330.63	390.95
TFT	bf_IF	297.98	272.76	323.19	286.99	264.28	309.71	353.11	322.24	383.97
TFT	if_IF	292.16	267.32	317.00	290.61	268.67	312.55	347.67	317.10	378.24
GD	BF	244.18	225.33	263.03	239.83	223.86	255.80	271.25	250.32	292.17
GD	bf_BF	242.76	223.88	261.64	237.50	221.48	253.52	267.48	246.51	288.44
GD	if_BF	242.03	222.32	261.75	238.61	221.81	255.41	261.34	239.47	283.21
GD	IF	259.63	239.20	280.06	250.22	232.30	268.13	297.03	271.12	322.94
GD	bf_IF	258.70	237.90	279.50	244.59	225.96	263.21	302.61	276.28	328.94
GD	if_IF	254.02	233.46	274.58	249.73	231.62	267.84	292.59	266.42	318.75
FFD	BF	215.13	200.63	229.63	209.39	199.50	219.27	221.60	212.16	231.05
FFD	bf_BF	215.00	200.47	229.52	208.53	198.61	218.46	221.42	211.93	230.91
FFD	if_BF	214.23	199.17	229.30	208.05	197.49	218.60	221.64	211.25	232.04
FFD	IF	225.20	210.03	240.37	212.10	202.08	222.12	230.15	219.04	241.26
FFD	bf_IF	229.08	213.67	244.48	212.00	201.41	222.60	234.60	223.04	246.15
FFD	if_IF	223.49	208.23	238.75	210.62	200.44	220.80	229.99	218.61	241.37
SKIP	BF	0.30	0.26	0.34	0.25	0.21	0.29	0.23	0.18	0.28
SKIP	bf_BF	0.30	0.26	0.35	0.26	0.22	0.30	0.24	0.19	0.29

SKIP	if_BF	0.29	0.24	0.34	0.28	0.23	0.33	0.22	0.17	0.27
SKIP	IF	0.20	0.16	0.25	0.17	0.14	0.22	0.17	0.13	0.23
SKIP	bf_IF	0.19	0.15	0.23	0.20	0.16	0.25	0.18	0.13	0.24
SKIP	if_IF	0.19	0.15	0.24	0.21	0.16	0.25	0.18	0.14	0.24
REG	BF	0.17	0.13	0.22	0.15	0.12	0.18	0.23	0.20	0.27
REG	bf_BF	0.16	0.12	0.20	0.13	0.11	0.16	0.22	0.18	0.26
REG	if_BF	0.16	0.12	0.21	0.16	0.13	0.19	0.21	0.17	0.26
REG	IF	0.17	0.12	0.23	0.19	0.16	0.22	0.23	0.18	0.28
REG	bf_IF	0.14	0.10	0.19	0.16	0.13	0.19	0.20	0.16	0.26
REG	if_IF	0.14	0.10	0.19	0.16	0.13	0.18	0.20	0.15	0.25

CHAPTER 3

The Role of Morphological Priming in Orthographic and Semantic Word Learning

This study has been submitted and currently under review in *Language Learning and Development* as Coskun, M., Kuperman, V. (submitted). The Role of Morphological Priming in Orthographic and Semantic Word Learning

Abstract

Understanding the morphemic structure of a word is crucial for expanding vocabulary knowledge. We examined the effect of morphological and identity priming on incidental word learning in adult L1 speakers. In the learning phase, low-frequency base words (e.g., caltrop) and novel derived words formed from those bases (e.g., caltroper) were embedded in short texts. We further administered post-tests of orthographic and semantic learning for 1) derived forms of the word, primed by two repetitions of the base form (e.g., caltrop-caltrop-CALTROPER), and 2) base forms preceded by two repetitions of the derived form (e.g., caltroper-caltroper-CALTROP), and 3-4) base or derived words primed by identical forms (e.g., caltrop-caltrop-CALTROP, caltroper-caltroper-CALTROPER). We found that identical priming led to optimal orthographic learning. Yet semantic learning of base forms (2, 3) was stronger, regardless of the priming type. We discuss reasons for the asymmetric effect of morphological structure on orthographic and semantic learning.

Introduction

Words with more than one morpheme form the majority in many languages, including English (Libben, Goral, & Baayen, 2017). Many of morphologically complex words are of low frequency and potentially unfamiliar to many language speakers (Baayen, 1994; Nagy & Anderson, 1984). Also, everyday dozens of new words entering the language are formed by different combinations of existing morphemes (70% of the words listed in an English dictionary, e.g., bioweapon, arborist, therapize; Algeo & Algeo, 1991). Therefore, the ability to recognize the internal structure of a word (morphological awareness), to infer the meaning of a word from the meanings of its morphological constituents (morphological processing) is crucial to the ability to acquire and produce new words (see among many others, Carlisle, 2000; Mahony, Singson, & Mann, 2000). Indeed, the role of morphology in the learning of language, reading, and vocabulary is a focus of highly prolific research in language instruction, educational psychology, applied linguistics, and psychology of language (Carlisle, 2003; Carlisle & Stone, 2005). Across studies, languages, age groups and experimental manipulations, higher indices of morphological awareness or processing ability come with better oral and written language skills (see Amenta & Crepaldi, 2012, for a review; also see Anglin, 1993; Carlisle, 2010; Clark & Cohen, 1984; Clark & Hecht, 1982; Deacon & Kirby, 2004; Droop & Verhoeven, 2003; Kuo & Anderson, 2006). Moreover, meta-analyses of intervention studies on English school-age children show that specific morphological instruction confers consistent moderate beneficial effects on their literacy (Bowers, Kirby, & Deacon, 2010; Goodwin & Ahn, 2010; 2013).

Existing theories of morphological processing vary widely in their assumptions about what morphological representations are is and how morphologically complex words are learned, stored in the mental lexicon, and retrieved during language production or comprehension (Baayen et al., 1997; Baayen, Milin, Đurđević, Hendrix, & Marelli, 2011; Frauenfelder & Schreuder, 1992; Giraudo & Grainger, 2001; Harm & Seidenberg, 1999, 2004; Plaut & Gonnerman, 2000; Schreuder & Baayen, 1995; Seidenberg & Gonnerman, 2000; Taft, 2004). Yet, to the extent that these theories concern themselves with learning, they tend to agree that the key to it is the existence of systematic co-occurrences of specific chunks of letters and sounds that effect a regular change in lexical meaning when added to words (Baayen et al., 2011; Kirby & Bowers, 2017; Perfetti, 2007; Rastle, 2019; Ulicheva, Harvey, Aronoff, & Rastle, 2020). For instance, a learner of written English can benefit from the regular co-occurrence pattern of the word-final letter chunk -ist, sound chunk /Ist/, and the facet of the word meaning that "denote[s] a person who practices or is concerned with something, or holds certain principles" (https://www.dictionary.com/browse/-ist). When the learned mapping between letters, sounds, and meanings of a morpheme is stable and "crisp", it can underlie at least two mechanisms of vocabulary growth, which we illustrate using the example above. One is *generalization*, or the ability to infer the (unfamiliar) meaning of a derived word (*X-ist*) from the pre-existing knowledge of the base *X* and of the meaning and function of the derivational morpheme (*-ist*). We will label another *decomposition*, or the ability to infer the (unfamiliar) meaning of the base *X* given the knowledge of the complex word *X-ist* and the derivational morpheme *-ist*. The paper asks whether these two learning mechanisms lead to equal learning outcomes or whether one of them is more successful than the other and why. We pursue this question in the novel word learning task that uses unfamiliar morphologically simple bases and their derivations (*caltrop* “a spiked weapon against cavalry” vs *caltroper* “a person who uses a spiked weapon against cavalry”). In the remainder of the Introduction, we review the relevant literature on lexical development, morphological learning and processing and outline the logic and design of the present study.

Morphology knowledge develops starting from the very early ages and has a crucial role in vocabulary growth (Anglin, 1993; Brysbaert, Stevens, Mandera, & Keuleers, 2016; Clark & Cohen, 1984; Clark & Hecht, 1982). As early as in the pre-literacy period, children use their morphological analysis skills to infer the meaning of morphologically complex words. For example, Clark and Cohen (1984) showed that 4- and 5-years old can successfully infer the meaning of the unfamiliar word *lifter* by relying on prior knowledge of the base morpheme *lift* and the agentive suffix *-er*. Similarly, they can also construct the relevant complex form from a given base morpheme to obtain the desired meaning. For example, children 3- to 5-years old successfully compose the complex form of a base form using agentive suffixes like *-ist* and *-er* (e.g., from *pov* to *pover*) to indicate agency (e.g., Clark & Cohen, 1984; Clark & Hecht, 1982). Multiple additional studies advocate the ability to use morphological knowledge (about the form, function, and meaning of morphemes) to generalize a base to the entire derivational family as an essential developmental step towards acquiring literacy and vocabulary growth (see Bertram, Laine, & Virkkala, 2000; Carlisle, 2000; Mahony et al., 2000 ; Wood, Mustian, & Cooke, 2012; Wysocki & Jenkins, 1987; and references above).

Surprisingly little is known, however, about the ability to infer the meaning of the unknown base word relying on the knowledge of a derived form and a derivational morpheme, the ability we labeled decomposition. Admittedly, it may be more common for a language learner to encounter the base word first rather than a derivation because simplex words tend to be more frequent and have a lower age-of-acquisition (Brysbaert & New, 2009; Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). Yet it is at least logically possible that some derived words may be encountered prior to their bases. Equipped with knowledge of the complex meaning, coupled with the knowledge of derivational morphology, a speaker may arrive at an inference regarding the meaning of the unseen base word. Anecdotally, we find it plausible that children may first acquire some facets of meanings about transparent words like “neurologist”, “plunger” or “(dental) retainer” before learning what “neurology”, “plunge” or “retain” means. Also, as argued in Nagy and Anderson (1984, see also Dawson, Rastle, & Ricketts, 2021), the proportion of low-frequency words with familiar affixes and unfamiliar base morphemes is substantial and ever-growing in the reading materials that adolescents and adults tend to encounter. In these cases, semantic decomposition – inferring the unknown base meaning from the derived form – may be commonly seen at work as a learning mechanism. Our main interest is in comparing mechanisms of generalization and decomposition as to how efficient they are for novel word learning, see below.

To our knowledge, very few studies have directly examined decomposition as a learning mechanism (though a vast literature has been dedicated to decomposition in morphological processing as the ability to segment a complex or pseudo-complex word into morphemes). One exception is Dawson et al. (2021) who investigated the contribution of affixes in semantic and orthographic word learning in English-speaking adolescents. They embedded non-word base morphemes with existing suffixes (e.g., *clantist*) into sentence contexts and then presented the target words with definitions matching or not matching the semantic and syntactic role of the suffix (the congruent vs incongruent condition). Post-tests showed stronger semantic learning of suffixed forms when their definition was congruent with the regular syntactic and semantic function of the suffix, but no parallel orthographic advantage was observed. The quantity that would shed light on the efficiency of decomposition as a learning mechanism would be the readers' acquired knowledge of the base form *clant*, which Dawson et al. do not report.

Decomposition as a learning mechanism was also the subject of a few studies interested in the orthographic learning. For example, Pacton, Foulin, Casalis, and Treiman (2013) investigated whether learners would benefit from morphological information to learn the spellings of novel words in orthographically opaque French. Third graders silently read short stories. In one condition, stories included seven repetitions of novel base morphemes ending with a final silent consonant (e.g., *vensois*), in the other condition, stories included five repetitions of those base morphemes and two repetitions of their derived forms (e.g., *vensoisist*), in which the base's final consonant was pronounced. Readers more successfully learned the spelling of a base morpheme (*vensois*) when they had a chance to read both forms (*vensois* and *vensoisist*) compared to when they only read the base form (*vensois*). The above result suggests that in certain situations exposure to the complex form of a word is more beneficial for learning the spelling of the base than exposure to the base itself, which can be interpreted as the advantage of decomposition. Another example comes from Ginestet, Shadbolt, Tucker, Bosse, and Deacon (2021). In their study, English-speaking adults read either morphologically complex nonwords (e.g., *re-lerb-er* or *mis-doaf-er*, the combination of real affixes with a nonword base) or monomorphemic orthographic controls with matching length (e.g., *pe-lerb-le* or *fis-doaf-le*) in short stories while their eye movements were recorded. The analysis of eye-tracking measures and the post-test of orthographic learning revealed that reading morphologically complex words resulted in faster and less effortful reading as well as faster and more accurate recognition of the spellings compared to the monomorphemic controls. The reported processing and learning advantage suggest that the segmentation of the morphologically complex words into morphemes occurred, which reduced the length of word to be learned (from the whole word to the base form) and decreased the quantity of new information that a reader needs to deal with.

Tucker, Castles, Laroche, & Deacon (2016), on the other hand, investigated both generalization and decomposition scenarios in learning the spellings of novel words. In their study, children were asked to read short stories including novel words from one of three conditions (e.g., *feap* for the base condition, *feaple* for the orthographically related condition, and *feaper* for the morphologically related condition). Then children performed an orthographic choice test including not only the novel words they were trained on (e.g., *feap,* if they had seen the word in the base condition), but also the alternate forms from the other two conditions (e.g., *feaple* and *feaper,* as orthographically and morphologically related forms of *feap*). Their task was to choose either the spelling they saw in the stories or the form orthographically or morphologically related to the seen form among their homophone distractors. The results showed that learners were able to extend their learning of the novel words seen in the stories to the orthographically and morphologically related unseen forms regardless of the direction of learning. This suggests that prior experience with a word will likely be an advantage when evaluating the spelling of other orthographically or morphologically related words, yet these findings do not necessarily suggest morphological decomposition or generalization as an underlying learning mechanism.

In sum, much more is known about the development and importance of morphological generalization –learning of the unknown complex word's meaning from the meanings of the base word and derivational morpheme – than about decomposition as a mechanism of learning an unknown meaning of the base from the complex word. This asymmetry highlights an undesirable lacuna both in theories of morphological processing and language learning and in applied work on morphological instruction and vocabulary acquisition. The present study is set up to compare learning gains of these two scenarios using the incidental word learning paradigm with repetition priming. We used the incidental learning paradigm since it has been associated with a high volume of vocabulary learning at older age groups, too (Anderson & Freebody, 1981; Nation, 2001). As a characteristic of the incidental word learning paradigm, participants were not aware that they will be tested on the novel words presented in short stories (see Hulstijn, 2003).

The key manipulation of the study was that we created prime-target pairs from low-frequency (and thus likely unfamiliar) base forms (e.g., *caltrop*, a spiky weapon used against infantry and cavalry) and the novel derived forms of these bases (e.g., *caltroper*). All prime and target items were embedded into short informative stories. In the learning phase, participants read the stories for comprehension and in the subsequent testing phase administered post-tests of orthographic and semantic learning for target words. The nature of the prime (derived or base word) and the target (derived or base word) determined the four experimental conditions of the learning phase. Specifically, the target words in these conditions were 1) derived forms of the word, primed by two repetitions of the base form (e.g., caltrop-caltrop-CALTROPER), 2) base forms primed by two repetitions of the derived form (e.g., *caltroper-caltroper-CALTROP*), and 3-4) words primed by identical base or derived forms (e.g., *caltrop-caltrop-CALTROP, caltroper-caltroper-CALTROPER*). In each condition, all instances of primes and targets appeared in a separate sentence. Condition (1) exemplified generalization in which the learned base word influences learning of the derived word with a familiar suffix, while condition (2) exemplified decomposition, i.e., learning of the derived word with a familiar suffix influenced the learning of the base word. These instantiations of morphological priming were complemented by the control conditions (3-4) of identity priming where the primes and the targets were identical (either both base words or both derived words). To our knowledge, this is the first systematic comparison of learning gains between the generalization and decomposition scenarios, aimed at determining the relative efficiency of each learning mechanism.

Priming is an extremely well-established experimental manipulation in morphological research (Schmidtke & Kuperman, 2020). Yet, as far as we know, only a few studies directly compared the amount of priming elicited in the base-derived and derived-base prime-target pairs. For instance, the semantically transparent condition of a masked priming study by Marslen-Wilson et al. (2008) reported a stronger priming effect at the SOA of 48 ms when the derived word (*bravely*) primed the target base (*brave*) than vice versa (when *brave* primed *bravely*), 36 ms vs 21 ms. While obtained from recognition of known words rather than learning of unknown ones, this finding hints at a processing advantage that decomposition confers relative to generalization. We tested whether this advantage is retained, nullified or reversed in the novel word learning paradigm.

The faciliatory role of a prime would also depend on the extent to which the lexical integration of this word has occurred. Typically, the full integration of a recently learned word to a mental lexicon occurs after a certain amount of time, like a consolidation period or several exposures to the word (e.g., Bowers, Davis, & Hanley, 2005; Gaskell & Dumay, 2003). The initial stages of word learning involve familiarization, when a learner acquires factual knowledge about the new word, such as its spelling, pronunciation, and meaning (termed lexical configuration by Leach & Samuel, 2007). A rapid identification of a newly learned word after a brief training can be an example of the acquisition of word knowledge that may or may not signify yet of a robust lexical representation (Gaskel & Dumay, 2003; Leach & Samuel, 2007). The distinguishing feature of a fully learned new word is linked with its capacity to interact with other words in the mental lexicon (termed lexical engagement by Leach and Samuel, 2007). As an outcome of lexical integration, a newly learned word can inhibit the recognition of the orthographically similar words (e.g. *cathedruke – cathedral*, Gaskell, & Dumay, 2003; *banara – banana*, Bowers, Davis, & Hanley; 2005) or support the activation of the semantically similar words (*fecton – cat*, in which *fecton* corresponds to the meaning of “a type of cat” Tamminen & Gaskell, 2013; also see Palma & Titone, 2021, for a review).

As a first demonstration of its kind, our study was deliberately designed to maximize the possible learning effects. First, we recruited adult proficient speakers of English, whom we may expect to have fully developed morphological awareness and knowledge. Second, we made use of highly productive English suffixes (*-ist* and *-er*) to form highly transparent derived forms, congruent with the typical meaning and syntactic function of those suffixes (animate nouns with agentive meaning). Third, we provided learning opportunities both for primes and targets by embedding each in a context specifically informative of the prime or the target. Finally, we doubled the occurrence of the prime to ensure the improved quality of its learning and the potential inference of this learning to either a morphologically simpler or a more complex form.

To sum up, our aim in this study is to investigate and compare the efficiency of decomposition and generalization as mechanisms of learning meanings of novel words. For this purpose, we create four conditions (see above), which instantiate all pairwise combinations of base and derived forms as prime and target. The efficiency of generalization and decomposition is determined by comparing the learning of target words in the conditions of morphological priming to respective identity priming conditions as control. We expect to find a difference in learning gains if one mechanism is more efficient than the other.

Methods

Participants

One hundred adult native English speakers participated in this experiment. We removed from consideration one participant, who provided the exact same fragments as they appeared in the reading material for the definitions of the words asked in the post-tests. The final sample consisted of 99 adults (age M=39.1, SD=12.5; age range 20 – 72, 52 self-identified as female, 46 as male, 1 as nonbinary). All participants were recruited through Amazon' Mechanical Turk (mturk.com) and compensated $3.70 for their participation. We determined this rate considering the duration of the experiment and the federal minimum wage, $7.25/hour in the U.S. The study was restricted to recruit native speakers of English residing in the U.S. Also, participants were asked to state their country of birth, their level of education, their languages known, and the order in which the languages were learned. This study was approved by McMaster University's Research Ethics Board (protocol 5020).

Materials

Target Stimuli

A total of sixteen words were used for the reading task, with eight existing base words and eight novel derived words which were formed from these bases using agentive suffixes *-ist* or *-er*. Base words corresponded to concrete nouns (e.g., *kinnor*, "a string instrument", *caltrop* "a weapon made of two or more sharp nails") and derived forms were intended to refer to occupations related to the base nouns (e.g., *kinnorist*, "a person who plays kinnor", *caltroper* "a person who uses caltrop", see also Supplementary Materials Table S1 for the full list of base words and their definitions.). Six of the base words were selected from the 1-billion-word Corpus of Contemporary American English (COCA) and an additional two from the 14-bilion-word iWeb Corpus. The Zipf-scale, calculated as $\log_{10}$(frequency per billion words), was used to estimate word frequencies in a comparable way across corpora. The mean word frequency for the selected words was 0.94 Zipf (SD = 0.83), with a range of 0-1.95. According to Van Heuven, Mandera, Keuleers, and Brysbaert (2014), words with the Zipf value below 3 are regarded as low frequency, thus all base words had a very low frequency of occurrence. While we cannot verify that our participants did not come even encounter the target words, it was safe to assume that for the majority of participants the majority of the words were unfamiliar. Derived forms were not attested in any of the corpora we inspected.

The experiment had a 2 x 2 design. One manipulation labeled Priming Type had two levels: identity priming (prime was identical to the target) or morphological priming (prime was morphologically related to the target). Another manipulation labeled Target Type determined the nature of the target, either a base form (e.g., *CALTROP*) or a derived form (e.g., *CALTROPER*). In this study, target words were always preceded by the two repetitions of the prime. We chose to provide two repetitions of the prime because it is likely to be the minimum number providing a critical level of form and semantic learning that would be detectable in our post-tests (see Hulme, Barsky, & Rodd, 2019; Ginestet, Valdois, Diard, & Bosse, 2020). Crossing the two factorial manipulations yielded four experimental conditions, represented in Table 1:

Table 1. The experimental conditions tested.

	Target Type: base	Target Type: derived
Priming Type: identity	caltrop-caltrop-CALTROP	caltroper-caltroper-CALTROPER
Priming Type: morphological	caltroper-caltroper-CALTROP	caltrop-caltrop-CALTROPER

We created eight prime-prime-TARGET sequences for each experimental cell presented in Table 1, to a total of 32 critical stimuli (see Supplementary Materials, Table S2 for the list of all stimuli).

Reading materials

Five-sentence paragraphs were written in the genre of narrative prose as contexts into which the prime and target words were embedded. The structure of the paragraphs was as follows: All paragraphs started with an introductory sentence and ended in a concluding sentence (as explained in the Target Stimuli). Neither of these sentences incorporated either of the primes or targets. Two identical instances of a prime were embedded into the second and third sentences, followed by a fourth sentence in which a target word was presented. None of the critical words was placed at the beginning or end of a line.

Table 2 illustrates the carrier paragraph for the four experimental conditions of the study: sequence [a-a-c] implements the identity priming + base target condition; sequence [b-b-d] identity priming + derived target condition; sequence [b-b-c] the morphological priming + base target condition; and sequence [a-a-d] the morphological priming + derived target condition.

Table 2. Example paragraphs. Sentences differing between conditions are italicised. Prime and target words are given in bold.

1st sentence	There is a deadly weapon that was used by armies in the past and up to the present day.
2nd sentence (first occurrence of prime)	[a. *The original **caltrop** was simply a multi-pointed metal star arranged*]/ [b. *A **caltroper** threw this multi-pointed metal star*] so that when three nails were on the ground the fourth was always pointed upward.
3rd sentence (second occurrence of prime)	While they first appeared thousands of years ago to defeat horses and soldiers, a [a. ***caltrop***]/ [b. ***caltroper***] is still a threat for wheeled vehicles.
4th sentence (target occurrence)	Factors like being cheap and easy to produce, being easily portable, needing no care or preparation, and requiring no special skills or training make [c. *a **caltrop** still popular as a weapon]/* [d. *this weapon handy to a **caltroper***] today.
5th sentence	With all that said, you can become dangerous simply by learning great footwork and movement.

The sentences with alternate forms differed minimally. Where possible, we created the second version of a sentence simply by replacing a critical stimulus (a base or a derived form) with its alternative so that the sentences differed by only one word (e.g., A [***caltrop/caltroper***] is still a threat). Where not possible, we attempted to choose the discrepant words from the same syntactic category (e.g., *oval-shaped symbol **of***/*oval-shaped symbol **with***). In total, around 4% of words differed in each version of these sentences. We also tried to keep the complexity of sentence structures comparable. For example, if a sentence in one version was a passive sentence, this structure was carried out to the other version as well (e.g., [*that they made **was called** an amaut*/*at that time **was made** by an amautist*]) (also, see Supplementary Materials, Table S2 for the full list of the materials).

Two researchers naïve to the purposes of this experiment were asked to read the paragraphs and guess the intended meaning of the target words. The necessary adjustments were made when they found the paragraphs not easy to read or they failed to infer the meaning of the target words. As a result, we obtained a total of 32 paragraphs, one for each prime target word pair, with 8 in each condition (see also Target Stimuli above). The paragraphs in each condition were assigned to one of the four lists in a counterbalanced manner so that each list contained an equal number of paragraphs (2 paragraphs) from each condition and each list contained only one of the four versions of the paragraphs. Thus, each participant would complete one list of 8 paragraphs.

Apparatus, and Procedure

We developed a web-based experiment using HTML and Javascript. Participants were recruited via an online crowdsourcing platform Amazon Mechanical Turk and eligible participants accessed the experiment through web browsers. Participants were assigned one of the four lists (described above). Before starting the experiment, they were provided an online consent form. Next, they filled out the demographic and language background questionnaires. Participants started the reading section with two practice trials. Then, they read 8 experimental passages presented one at a time on the screen in random order. Post-tests of orthographic decision, definition prompting, and definition matching tests were administered in this order after the reading section. The online debrief form was presented at the end of the study and the entire experiment took around half an hour.

Vocabulary post-tests: The following three tests were administered to measure the acquired form and meaning knowledge of the target words presented in the texts: Orthographic decision, definition matching, and definition prompting tasks. Only the word knowledge for target words encountered during the passage reading was measured.

The *Orthographic Decision Task* was designed to assess orthographic learning of the target words. A homophonic nonword alternative was generated for each target word (e.g., *caltrop* as word – *kaltrop* as foil). In this forced choice test, participants were asked to decide as quickly and accurately as possible whether they had encountered the presented word during reading. Items were presented on the screen one at a time, in a random order. No feedback was given during the task. Both the binary index of accuracy and response latency were collected.

The *Definition Prompting Task* tapped into participants' ability to recall meaning of the target words. In this semantic recall task, participants were asked to provide the definition of the items presented on the screen. They had a chance to ask for up to two hints if they could not come up with a guess. The hints were segments of the texts they read containing this word. We operationalized the outcomes of this task in two ways. Under the first operationalization, the correct responses on the first attempt without taking a hint were scored three points. The total score reduced one point after each hint and had a minimum of zero, if a correct definition was not provided after 2 hints. The outcome was a score in the 0—3 range, and the baseline achieved if responses were given at random was 0 points. A second outcome of the test additionally examined whether the given answers were related to the alternative form of the target word. For example, one might come up with a definition matching with the word *kinnor* (an instrument) for the question asking the definition of the word *kinnorist* (a performer). We calculated the proportion of these answers to investigate whether there is any tendency to recall the meaning of a morphologically related form better than that of the actual form.

The *Definition Matching Task* assessed the ability to recognize the meaning of the target words. The target words were presented along with the eight unseen foil words on one side of the screen, and the definitions of the target words were presented along with the eight foil definitions on the other side. Foil words were created from the target words using Wuggy (Keuleers & Brysbaert, 2010), maintaining letter length, number of syllables, and two out of three of the sub-syllabic segments. The presented words and definitions appeared in a random order. Participants were instructed to match the words they read earlier with the best-fitting definition. The outcome was a binary accuracy measure. The baseline performance in this test, which would be achieved if the targets and definitions were matched randomly, is 6.2%.

Variables

Accuracy scores from all post-tests and response times (RTs) from the orthographic decision test were analyzed as dependent variables.

The two critical independent variables were binary factors Priming Type and Target Type (see Table 1): we examined their main effects and interaction on the learning outcomes. Since the total time spent on each text might tap into the amount of attention given to each text, logged total passage reading time was also introduced as a control variable. We also considered the trial number in orthographic decision and definition prompting tests as a control variable as long as it did not lead to any convergence issues in the model.

We fitted generalized linear mixed effect regression models (using the lme4 package, Bates, Maechler, Bolker, & Walker, 2015, in the statistical software environment R, R Core Team, 2018) to the dependent variables described above using independent variables Priming Type and Target Type, as well as by-subject and by-word random intercepts. We fitted models with the Poisson family to count variables (e.g., test scores ranging from 0 to 3), and the binomial family to binary variables. As recently recommended by Lo and Andrews (2015), the Gamma family and the identity link were used for reaction times (RTs) data to account the right-skewed shape of the distribution without any need for data transformation. Recent studies suggest RT measures collected from the web- and lab-based studies offer comparable results (e.g., de Leeuw & Motz, 2016; Anwyl-Irvine, Dalmaijer, Hodges, & Evershed, 2020).

Results

Orthographic decision test

We used median absolute deviation (MAD) around the median to determine outliers (Leys, Ley, Klein, Bernard, & Licata, 2013). Trials with response times more than 3.5 MAD from the median were excluded from the analysis (17%), (M_{RT} = 1690, SD = 792, RT range 332 – 4585 ms). Although this might seem like a higher proportion of the data, we should note that around 12% of the responses exceed 10 s response time. Since orthographic choice decisions are typically made within the first few seconds after an item is presented (< 2 s, see, for example, Ginestet et al., 2020, 2021), the response times as long as 10 s are more likely to indicate a participant's lack of attention to the task, rather than their decision-making effort.

The baseline for this test, which would be achieved if either answer was selected randomly, is 50%. Participants showed an above-chance level of form learning (M=0.70, SD=0.46). Accuracy in acceptance of target words and in correct rejection of the foil words were significantly above the baseline, as indicated in the proportion test against the 50% baseline, $p < 0.01$). We excluded the foils from the rest of the analysis and performed RT and accuracy analyses for the responses to target words. We first analyzed accuracy of orthographic decision test responses (659 responses) as a function of Priming Type and Target Type. Figure 1 (left) shows that the identical repetition priming has a stronger effect on accuracy scores in orthographic decision task than the morphological priming (68% morphological vs 81% identical, see Supplementary Materials, Table S3 for descriptive statistics). This main effect is confirmed in the mixed-effects logistic regression model reported in Supplementary Materials, Table S4. Regardless of the form of the target words (base or derived), the form was learned better when the repeated words were identical rather than morphologically related. Thus, neither generalization nor semantic decomposition reach the same efficiency in orthographic learning as pure repetition.

Then, we analyzed RTs to seen and correctly identified target words (491 responses). Figure 1 (right) summarizes these results and shows that participants responded faster to base target forms than to derived ones (M = 1496 ms, SD = 644 for base forms, and M = 1651 ms, SD = 779 for derived forms). Also, RTs to correctly identified base forms were similar regardless of the form of the prime preceding the target word. However, the recognition of the derived target form accelerated only when the prime and target words were identical. In sum, there was a speed advantage to orthographic choice response times in the condition that we associate with semantic decomposition (derived-derived-BASE) as compared to the generalization condition (base-base-DERIVED). The former was numerically indistinguishable from the identity priming condition (base-base-BASE). The main effect of Target Type and the interaction effect between Target Type and Priming Type were also confirmed by the models, reported in Supplementary Materials, Table S4.

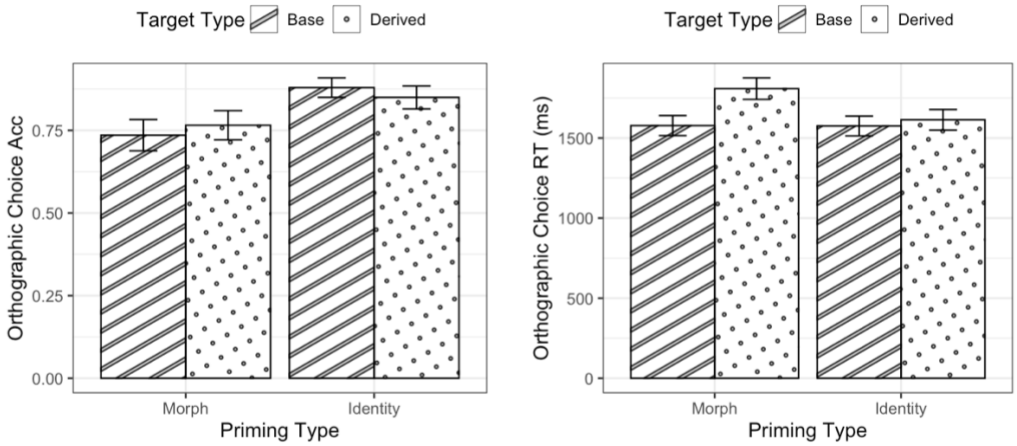

Figure 1. Partial effects of Priming Type and Target Type on orthographic decision accuracy (left) and RT (right). Error bars show ±1 standard error.

Definition Matching Test

Figure 2 (left) shows the effects of Priming Type and Target Type on definition matching accuracy. Participants performed equally well across the conditions, with accuracy around 45% in each, see Supplementary Materials, Table S3. As indicated by the post-hoc proportion test, responses were more accurate than the baseline of 6.2% accuracy ($p < 0.01$). Neither Priming Type nor Target Type did influence semantic recognition performance, which was also confirmed by the regression model (see Supplementary Materials, Table S4)

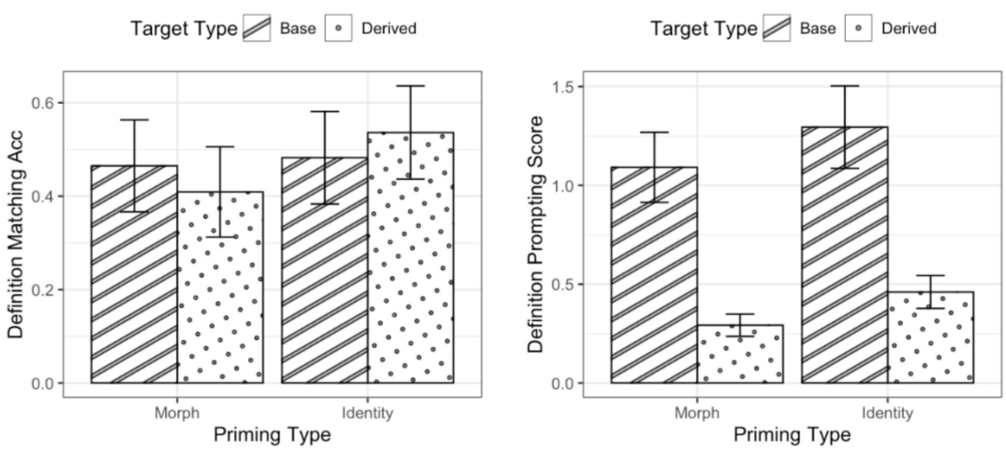

Figure 2. Partial effects of Priming Type and Target Type on definition matching scores (left) and definition prompting scores (right). Error bars show ±1 standard error.

Definition Prompting Test

A total of 792 responses to seen target words were collected. Participants showed an above-chance level performance in recalling the meanings of the target words (M=0.96, SD = 1.24). Figure 2 (right) summarizes the estimated definition prompting scores by Target Type and Priming Type. Participants were much more likely to recall the meaning of a target word when it was a base form than when it was a derived form (M = 1.4, SD = 1.3 for the base forms, M= 0.5, SD = 1.0 for the derived forms). Figure 2 (right) also shows that the repetition of the same form led to a better performance in this semantic recall task than the repetition of the morphologically related form, (M = 1.07, SD = 1.3 for the identical repetition vs M= 0.84, SD = 1.2 for the morphological repetition). In other words, morphological priming of the base/derived word led to semantic learning of the same magnitude as identity priming of the base/derived word, with a strong advantage of learning the meaning of the base rather than the derivation. These main effects were also confirmed in the mixed-effects model reported in Supplementary Materials, Table S4. There was no interaction between Priming Type and Target Type.

A closer inspection of the actual definitions typed in response to different types of stimuli added nuance to this picture. There were 396 responses in each condition of the Target Type (base vs derived) collapsed over identity and morphological Priming Types. In their responses to the base forms, participants recalled the correct definition of those forms 62% of the time, whereas this percent was much lower (21%) for the derived forms. This imbalance arose because participants often produced a definition for the base form when asked to define the derived word they learned (e.g., with *caltroper* as a prompt, they would respond with “a weapon” instead of “a soldier”, where “a weapon” is a definition of *caltrop*). This substitution of the definition when base was defined instead of derived word took place 37% of the time when derived words were a prompt to the task.). The opposite was nearly never true: participants only provided a derived form definition when responding to the base form 1% of the time. Table 3 summarizes the distribution of the frequencies of the participants’ answers based on which form of the word was recalled and which form had been asked in the definition prompting test. The asymmetry in the direction of substitutions as a function of Target Type (and collapsed over Priming Type) was confirmed by the chi-square test (X^2 (1, N = 479) = 89.64, p <0.05). This finding shows strong support for semantic decomposition. Participants tend to decompose the derived target word and produce the definition for the base instead more often than they produce the definition for the actual derived prompt (37% vs 21%). (Over)generalization, or producing the derived word definition when prompted with the base stimulus, is exceedingly rare (5%).

Table 3. The distribution of the frequencies of the given answers based on the Target Type of the recalled word form (identity and morphological priming conditions jointly)

	The answer provided by participants	
	Base	Derived
Target Type in the paragraph: Base	245/396 (62%)	5/396 (1%)
Derived	147/396 (37%)	82/396 (21%)

General Discussion

Morphological awareness and knowledge give rise to two possible learning mechanisms, generalization (from known base to unknown derived) and decomposition (from known derived to unknown base). The role of generalization is very well attested (e.g., Bertram, Laine, & Virkkala, 2000; Carlisle, 2000; Mahony et al., 2000; Wood, Mustian, & Cooke, 2012; Wysocki & Jenkins, 1987), while the role of decomposition much less so (Dawson et al., 2021; Ginestet et al., 2021; Pacton et al., 2013; Tucker et al., 2016). Possibly this is because early years offer more opportunities for generalization. Yet more advanced reading materials give rise to more occasions when readers come across derived forms first and need to infer the meaning of the base (e.g., Nagy & Anderson, 1984). Because of the relevance for theories of morphological processing and word learning, we examined the learning mechanisms of generalization and decomposition directly. Our experiment compared the two learning mechanisms as morphological priming in an implicit novel word learning paradigm, paired up with identical priming as control.

For the purpose of this study, we presented L1 adult speakers short stories containing all pairwise combinations of base and derived forms as prime and target (see the Introduction). The prime-target set represented by the base-derived words exemplified generalization, the derived-based words exemplified decomposition, and base-base and derived-derived identical sets served as controls. We examined the orthographic and semantic learning of target words to evaluate the efficiency of generalization and decomposition. If one mechanism is more efficient than the other, we expected to find a difference in learning gains.

The results revealed two noteworthy points. First, identity priming was always as good or better than morphological priming. The accuracy of semantic and orthographic learning outcomes, and the speed of recognition of the correct spelling of the target word showed that target words in the identity priming condition were learned better and required less effort to be retrieved. Taken together, these results suggest that the repetition of the precisely same form would be more likely to increase the quality of learning. Thus, in instructional materials, the learning gains for morphologically simple or complex words are the highest if more exposure is added to the same form rather than morphologically related forms.

Second, semantic decomposition appears to be a stronger, more efficient mechanism of both the orthographic and semantic learning. Participants were faster to respond to the base in the orthographic choice task in the "decomposition" (derived-derived-BASE) condition, where most learning about the base came through the derived word. Also, the analyses of errors in the definition prompting task revealed that participants defaulted to semantic decomposition when prompted for definitions. With a derived word as a stimulus to define in the task, they would more often revert to the meaning of the base rather than the actual derived word. Our results are consistent with the findings of previous studies which reported the advantage of decomposition in the learning of the spelling of the new words (Pacton et al., 2013; Ginestet et al., 2013).

One possible explanation for why generalization is not as efficient as decomposition as a learning mechanism could be related to the difference in the amount of information provided by base and derived forms. The meaning of a base form would naturally exist in the meaning of its semantically transparent complex form (e.g., a *caltoper* is a soldier who uses a *caltrop* as a weapon). Therefore, in the decomposition scenario, each occurrence of a derived form would also create an opportunity for a learner to gain information about the meaning of the base which facilitates to infer its meaning. In the generalization scenario, however, a base form does not provide any direct clue about the form or meaning of its potential derived form. As a result, the necessary information to infer the meaning of the derived form cannot be available in advance. With the occurrence of a derived form, a learner must identify both the meaning of a base form and the role of the suffix in the word to be able to successfully infer the meaning of the complex form. Compared to the same process for a simplex form, this requires more effort and as a result, might be the underlying reason leading to the inefficiency of generalization as a learning mechanism. Taken together, these findings suggest that learning the meanings of simple words through an inference from complex words and morphological knowledge may be a valuable and under-used instructional tool.

We set up the study in the way that would potentially highlight the effects (adult proficient readers, highly productive and orthographically and semantically transparent suffixes, two repetitions of the prime, contextual information given on both the prime and the target). This line of research needs to be extended over less proficient readers and less clear-cut morphological contexts. Since less proficient readers may be less experienced with learning new words through decomposition, the observed advantage of decomposition may not emerge in that group. Similarly, less transparent morphologically complex forms (e.g., depart-department) may not grant a learning advantage as the semantic relationship between these lexical items is less noticeable, which may make decomposition more difficult.

This study demonstrates the importance of a systematic comparison of possible learning mechanisms – decomposition and generalization – for understanding the relationship between morphology and orthographic and semantics aspects of word learning. The outcomes of this study can also help educators develop efficient word-learning strategies.

References

Algeo, J., & Algeo, A. (1991). Among the new words. *American Speech*, *66*(4), 380-406.

Amenta, S., & Crepaldi, D., 2012. Morphological Processing as We Know It: An Analytical Review of Morphological Effects in Visual Word Identification. *Frontiers in Psychology*, 3.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), Comprehension and Teaching: Research Reviews (pp. 77-117). Newark, DE: International Reading Association.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society of Research in Child Development, 58*(10, Serial No. 238).

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*, *53*(4), 1407-1425.

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

Baayen, R. H. (1994). Productivity in language production. *Language and cognitive processes*, *9*(3), 447-469.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological review*, *118*(3), 438.

Bertram, R., Laine, M., & Virkkala, M. M. (2000). The role of derivational morphology in vocabulary acquisition: Get by with a little help from my morpheme friends. *Scandinavian Journal of Psychology*, 41(4), 287-296.

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Interfering neighbours: The impact of novel word learning on the identification of visually similar words. *Cognition*, *97*(3), B45-B54.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of educational research*, *80*(2), 144-179.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, *41*(4), 977-990.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E., 2016. How Many Words Do We Know? Practical Estimates of Vocabulary Size Dependent on Word Definition, the Degree of Language Input and the Participant's Age. *Frontiers in Psychology*, 7.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and writing*, *12*(3), 169-190.

Carlisle, J. F. (2003). Morphology matters in learning to read: A commentary. *Reading Psychology*, *24*(3-4), 291-322.

Carlisle, J. F., & Stone, C. A. (2005). Exploring the role of morphemes in word reading. *Reading research quarterly*, *40*(4), 428-449.

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading research quarterly*, *45*(4), 464-487.

Clark, E. V., & Cohen, S. R. (1984). Productivity and memory for newly formed words. *Journal of Child Language*, *11*(3), 611-625.

Clark, E. V., & Hecht, B. F. (1982). Learning to coin agent and instrument nouns. *Cognition*, *12*(1), 1-24.

Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness: Just "more phonological"? The roles of morphological and phonological awareness in reading development. *Applied psycholinguistics*, *25*(2), 223-238.

Dawson, N., Rastle, K., & Ricketts, J., 2021. Bridging form and meaning: support from derivational suffixes in word learning. *Journal of Research in Reading*, 44(1), pp.27-50.

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, *48*(1), 1-12.

Droop, M., & Verhoeven, L., 2003. Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, 38(1), pp.78-103.

Frauenfelder, U. H., & Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In *Yearbook of morphology 1991* (pp. 165-183). Springer, Dordrecht.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*(2), 105-132.

Ginestet, E., Valdois, S., Diard, J., & Bosse, M. L. (2020). Orthographic learning of novel words in adults: effects of exposure and visual attention on eye movements. *Journal of Cognitive Psychology*, *32*(8), 785-804.

Ginestet, E., Shadbolt, J., Tucker, R., Bosse, M. L., & Deacon, S. H. (2021). Orthographic learning and transfer of complex words: insights from eye tracking during reading and learning tasks. *Journal of Research in Reading*, *44*(1), 51-69.

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review*, *8*(1), 127-131.

Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of dyslexia*, *60*(2), 183-208.

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257-285.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological review*, *106*(3), 491.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological review*, *111*(3), 662.

Hulme, R. C., Barsky, D., & Rodd, J. M. (2019). Incidental learning and long-term retention of new word meanings from stories: The effect of number of exposures. *Language Learning*, *69*(1), 18-43.

Hulstijn, J. (2003). 12 Incidental and Intentional Learning. *The handbook of second language acquisition*, *19*, 349.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior research methods*, *42*(3), 627-633.

Kirby, J. R., & Bowers, P. N. (2017). Morphological instruction and literacy. *Theories of reading development*, *15*, 437.

Kuo, L. J., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational psychologist*, *41*(3), 161-180.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. Behavior research methods, 44(4), 978-990.

Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive psychology*, *55*(4), 306-353.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of experimental social psychology*, *49*(4), 764-766.

Libben, G., Goral, M., & Baayen, H. (2017). Morphological integration and the bilingual lexicon. *Bilingualism: A framework for understanding the mental lexicon*, *6*, 197.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in psychology*, *6*, 1171.

Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and writing*, *12*(3), 191-218.

Marslen-Wilson, W., Bozic, M., & Randall, B., 2008. Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes*, 23(3), pp.394-421.

Nagy, W., & Anderson, R., 1984. How Many Words Are There in Printed School English?. *Reading Research Quarterly*, 19(3), p.304.

Nation, I. S. (2001). *Learning vocabulary in another language*. Ernst Klett Sprachen.

Pacton, S., Foulin, J., Casalis, S., & Treiman, R., 2013. Children benefit from morphological relatedness when they learn to spell new words. *Frontiers in Psychology*, 4.

Palma, P., & Titone, D., 2020. Something old, something new: A review of the literature on sleep-related lexicalization of novel words in adults. *Psychonomic Bulletin & Review*, 28(1), pp.96-121.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading*, *11*(4), 357-383.

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing?. *Language and Cognitive Processes*, *15*(4-5), 445-485.

Rastle, K. (2019). The place of morphology in learning to read in English. *Cortex, 116*, 45-54.

Schmidtke, D., & Kuperman, V. (2020). Psycholinguistic Methods and Tasks in Morphology. In *Oxford Research Encyclopedia of Linguistics*.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. *Morphological aspects of language processing*, *2*, 257-294.

Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in cognitive sciences*, *4*(9), 353-361.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A*, *57*(4), 745-765.

Tamminen, J., & Gaskell, M. G. (2013). Novel word integration in the mental lexicon: Evidence from unmasked and masked semantic priming. *Quarterly Journal of Experimental Psychology*, *66*(5), 1001-1025.

Tucker, R., Castles, A., Laroche, A., & Deacon, S., 2016. The nature of orthographic learning in self-teaching: Testing the extent of transfer. *Journal of Experimental Child Psychology,* 145, pp.79-94.

Ulicheva, A., Harvey, H., Aronoff, M., & Rastle, K. (2020). Skilled readers' sensitivity to meaningful regularities in English writing. *Cognition*, *195*, 103810.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly journal of experimental psychology*, *67*(6), 1176-1190.

Wood, C. L., Mustian, A. L., & Cooke, N. L. (2012). Comparing whole-word and morphograph instruction during computer-assisted peer tutoring on students' acquisition and generalization of vocabulary. *Remedial and Special Education*, *33*(1), 39-47.

Wysocki, K., & Jenkins, J. R. (1987). Deriving word meanings through morphological generalization. *Reading Research Quarterly*, 66-81.

Supplementary Materials

Table S1. The list of base words with their definitions.

Word	Definition
amaut	is a kind of protective outer covering
bulla	is a kind of jewelry
caltrop	is a kind of weapon
cockade	is a kind of small accessory
gimbal	is a kind of tool
kalpis	is a kind of pot.
kinnor	is a kind of musical instrument.
tumbril	is a kind of vehicle

Table S2.The list of prime and target words used in the experiments, and the paragraphs in which these words were embedded. For demonstration purposes, the differences between the sentences with alternate forms of a word are italicized and the prime and target words in paragraphs are indicated with bold. Conditions—Identical Priming + Target Base: a + c, Identical Priming + Target Derived: b + d, Morphological Priming + Target Base: b + c, Morphological Priming + Target Derived: a + d

	1. Amaut/Amautist		
Prime	a. amaut - amaut b. amautist - amautist	1st sentence	Once upon a time, when there were not any developed tools, people used only their hands to make complex clothing.
		2nd sentence	One of the most important costumes [a *that they made was called an* ***amaut***]/ [b. *at that time was made by an* ***amautist***] and was specially designed for carrying children under the age of two.
		3rd sentence	[a. *The* ***amaut*** *was*]/ [b. *The* ***amautist*** *creates*] a parka with a large comfortable bag on the back just below the hood for babies.
Target	c. amaut d. amautist	4th sentence	[c. *The* ***amaut***]/ [d. *The design by the* ***amautist***] keeps the baby warm and safe from the wind and the cold, and also helps the mother and child to bond.

		5th sentence	Bonding at an early age helps them to develop a healthier relationship later on.
	2. Bulla/Bullar		
		1st sentence	In the Roman culture, children were seen as being very vulnerable and needing protection.
Prime	a. bulla - bulla b. bullar - bullar	2nd sentence	That is why newborn infants would wear a [a. ***bulla** which is a special kind of a necklace*]/ [b. *special kind of a necklace which is made by a **bullar***] to protect them from misfortune or injury.
		3rd sentence	Typically, [a. *a **bulla** was worn around the arm or the neck, sometimes with several of them together on one chain, and might contain perfume.*]/ [b. a ***bullar** fashioned this for the arm or the neck, sometimes put several of them together on one chain, and might add perfume to it*.]
Target	c. bulla d. bullar	4th sentence	[c. *A **bulla** could be]/* [d. *A **bullar** could make*] as simple as a knotted string of cheap leather or as elaborate as a finely made chain necklace holding a golden piece
		5th sentence	The practice was widespread throughout the Mediterranean and continued well beyond the ancient period.
	3. Caltrop/Caltoper		
		1st sentence	There is a deadly weapon that was used by armies in the past and up to the present day.
Prime	a. caltrop-caltrop b.caltroper-caltroper	2nd sentence	[a. *The original **caltrop** was simply a multi-pointed metal star arranged*]/ [b. *A **caltroper** threw this multi-pointed metal star*] so that when three nails were on the ground the fourth was always pointed upward.
		3rd sentence	While they first appeared thousands of years ago to defeat horses and soldiers, [a. *a **caltrop***]/ [b. *a **caltroper***] is still a threat for wheeled vehicle.

Target	c. caltrop d. caltroper	4th sentence	Factors like being cheap and easy to produce, being easily portable, needing no care or preparation, and requiring no special skills or training make [c. *a **caltrop** still popular as a weapon today]/* [d. *this weapon handy to a **caltroper** today*].
		5th sentence	With all that said, you can become dangerous simply by learning great footwork and movement.
	4. Cockade/Cockader		
		1st sentence	From the 15th century, coloured small accessories were used in Europe to show the class or rank of their wearers in a society.
Prime	a. cockade - cockade b. cockader - cockader	2nd sentence	Basically, [a. *a **cockade** is a circular or oval-shaped symbol of*]/ [b. *a **cockader** makes a circular or oval-shaped symbol with*] distinctive colours, which can be imagined as an imitation of a rose, and is usually made of fabric.
		3rd sentence	In the military, [a. *a **cockade** was worn in soldiers' hats to*]/ [b. *a **cockader** tailored these accessories for soldiers' hats to help*] distinguish members of opposing sides: the colors change depending upon the soldiers' nationality.
Target	c. cockade d. cockader	4th sentence	Today, a [c. ***cockade** is more of a fashionable object that]/* [d. ***cockader** creates a more fashionable version of this object which*] can add a lot of interest to a hat, be pinned to a jacket, and even be tied onto your shoes
		5th sentence	As you can imagine, this is not the only thing from the past that has become trendy today.
	5. Gimbal/Gimbalist		
		1st sentence	When filming a video, it is important to keep the camera as steady as possible.
Prime	a. gimbal-gimbal b. gimbalist - gimbalist	2nd sentence	A [a. ***gimbal***]/ [b. ***gimbalist***] can help you do this by keeping the camera horizontally stable all the time even if its supports shake
		3rd sentence	With [a. *the help of a **gimbal***]/

			[b. *help from a **gimbalist***] you can capture some smooth shots that would otherwise look very shaky.
Target	c. gimbal d. gimbalist	4th sentence	If you want to make your video look as high quality as a movie, make sure to [c. *use a **gimbal**]/* [d. *hire a **gimbalist***]. during your filming process
		5th sentence	All it takes is a few small improvements like this to bring your project to the next level.
	6. Kalpis/Kalpiser		
		1st sentence	The ancient Greeks used a variety of pottery in different shapes and sizes.
Prime	a. kalpis - kalpis b. kalpiser- kalpiser	2nd sentence	[a. *The **kalpis** was a container used for storing water*,]/ [b. *One of them was a container carried by a **kalpiser** to store water*,] which had two handles opposite each other at the level of the pot's shoulders.
		3rd sentence	These horizontal handles allowed [a. *the **kalpis** to be lifted, carried, and emptied*]/ [b. *the **kalpiser** to lift, carry, and empty it*] more easily
Target	c. kalpis d. kalpiser	4th sentence	There are a number of ancient Greek pictures that show [c. *a woman filling a **kalpis**]/* [d. *a **kalpiser** filling one*] at a fountain and balancing it on the top of her head to make it easier to carry
		5th sentence	Such scenes help us understand everyday life in ancient Greece.
	7. Kinnosr/Kinnorist		
		1st sentence	While we may never know for sure when human beings first developed music, we do know that some of the earliest examples appeared over 40,000 years ago.
Prime	a. kinnor-kinnor b.kinnorist- kinnorist	2nd sentence	One example is [a. a ***kinnor*** which was used]/ [b. a ***kinnorist*** who played] in the first centuries in Israel on family occasions and at popular festivals.
		3rd sentence	Local documents depict [a. *a **kinnor** as an instrument*]/ [b. a ***kinnorist*** with an instrument]

			that is made of wood and has ten strings
Target	c. kinnor d. kinnorist	4th sentence	In old books, David was depicted as [c. playing a ***kinnor***]/ [d. a ***kinnorist*** and as playing] to King Saul by holding it in his arms and using his fingers to pick the strings.
		5th sentence	To historians, this also shows a great friendship between two important people.
	8. Tumbril/Tumbriler		
		1st sentence	The ancient Sumerians lived thousands of years ago in a region in The Middle East.
Prime	a. tumbril- tumbril b. tumbriler-tumbriler	2nd sentence	The Sumerians didn't invent wheeled vehicles, but they developed the first two-wheeled [a. ***tumbril*** that was controlled by a person]/ [b. one controlled by a ***tumbriler***] and pulled by a donkey and other animals instead of horses.
		3rd sentence	In war, [a. *a **tumbril** often carries two standing people: a driver*]/ [b. *they often carry two standing people: a **tumbriler***] and a fighter using a bow-and-arrow.
Target	c. tumbril d. tumbriler	4th sentence	It's important that [c. *one should stand close to the wheels of a **tumbril***]/ [d. *a **tumbriler** should stand close to the wheels*] to have a better center of balance and for improved control over the vehicle.
		5th sentence	Although they were very common at early times, they have disappeared over the years.

Table S3. Summary of Orthographic Choice, Definition Matching, and Definition Prompting Tests Results

		Morphological Priming			Identity Priming		
	Target Type	Mean	SD	SE	Mean	SD	SE
Ortho							
RT	B	1616.04	705.05	56.09	1479.51	615.62	47.50
RT	D	1711.30	820.33	63.29	1604.12	752.07	58.55
Acc	B	0.66	0.47	0.04	0.83	0.38	0.03
Acc	D	0.70	0.46	0.04	0.79	0.41	0.03
Def Matching							
Acc	B	0.45	0.50	0.04	0.46	0.50	0.04
Acc	D	0.41	0.49	0.04	0.47	0.50	0.04
Def Prompt							
Score	B	1.29	1.24	0.09	1.53	1.29	0.09
Score	D	0.39	0.95	0.07	0.62	1.12	0.08

Table S4.Summary of the generalized linear mixed effect regression models fitted to the accuracy and RTs of orthographic choice task, the accuracy of definition matching task, and the definition prompting exact.

	Orthographic Choice Acc			**Orthographic Choice RT**			**Definition Matching Acc**			**Definition Prompting Scores**		
Fixed Effects	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*
Intercept	-1.67	1.18	0.159	1289.49	48.93	**<0.001**	-1.64	0.61	**0.007**	0.04	0.24	0.870
Target Type (derived)	0.16	0.30	0.591	230.85	32.11	**<0.001**	-0.23	0.44	0.606	-1.32	0.23	**<0.001**
Rep Type (identical)	0.96	0.28	**0.001**	-2.39	35.24	0.946	0.07	0.28	0.803	0.17	0.08	**0.042**
Trial Number	-0.08	0.02	**<0.001**	-25.97	4.90	**<0.001**				-0.02	0.01	**0.023**
Log Passage Reading Time	0.32	0.12	**0.007**	47.15	8.40	**<0.001**	0.41	0.15	**0.005**	0.05	0.05	0.299
Target Type (derived) x Rep Type (identical)	-0.41	0.39	0.287	-192.11	36.78	**<0.001**	0.44	0.40	0.269	0.28	0.17	0.091
Random Effects												
σ^2	3.29			0.14			3.29			0.93		
τ_{00}	0.49 $_{user\_id}$			83821.19 $_{user\_id}$			5.07 $_{user\_id}$			0.40 $_{user\_id}$		
	0.10 $_{target}$			3148.30 $_{target}$			0.46 $_{target}$			0.14 $_{target}$		
N	93 $_{user\_id}$			91 $_{user\_id}$			99 $_{user\_id}$			99 $_{user\_id}$		
	16 $_{target}$			16 $_{target}$			16 $_{target}$			16 $_{target}$		
Observations	659			491			792			792		
Marginal R^2 / Conditional R^2	0.086 / 0.224			0.225 / 1.000			0.021 / 0.634			0.208 / 0.497		

CHAPTER 4

Exposure and Attention in Novel Word Learning in the First and Second Language

This chapter has been submitted and currently under review *Journal of Cognitive Psychology* as Coskun, M., Lana, N., Kuperman, V. (submitted). Exposure and Attention in Novel Word Learning in the First and Second Language.

Abstract

This study investigated effects of exposure and selective attention on novel word learning in both native (L1) and non-native (L2) speakers of English. Attention to novel words was manipulated by delivering attention-inducing instructions, while the control group received no instructions. Exposure was manipulated by including a novel word either 2, 4, or 8 times throughout a passage. Participants read passages while having their eye movements tracked. Post-tests were performed immediately after and one week later. We found that the attention-inducing manipulation of instruction led to a short-lived speed-up in eye-movements and led to faster recognition of novel words. Critically, we observed that 4 exposures to a novel word led to the maximum post-test performance in L1 readers, which was not further improved through additional exposures. In L2 readers, however, learning continued to improve throughout the 8 exposures.

Introduction

Vocabulary acquisition is arguably the most challenging part of learning a language (Nation, 2001; Saragi, Nation, & Meister, 1978; Schmitt, 2008). Word learning can happen as a result of direct teaching, intentional learning (Nation, 2001), or without any learning intention as in incidental learning (Schmitt, 2008). While there are clear advantages to intentional learning, thus often being used in language teaching, the bulk of vocabulary acquisition is attributed to incidental learning (Anderson & Freebody, 1981; Nation, 2001). The important roles of both intentional and incidental learning in vocabulary acquisition in one's own language (L1) or a foreign language (L2) highlight the importance of understanding the factors that contribute to the quality of these learning types both within and across language speakers.

Two of the many dimensions proposed as being influential for incidental word learning are the amount of exposure to a novel word in a learner's input and selective attention that the learner directs to the word during those exposures. The present study examines these factors and an interaction between them in a novel-word learning paradigm among L1 and L2 speakers of English. Below we review relevant evidence from the prior word learning literature and present the rationale and the predictions of our experiment.

Exposure

The number of exposures to a novel word has a robust facilitatory effect on vocabulary learning (Mohamed, 2018; Horst, 2005; Pellicer-Sánchez, 2016; Webb, 2008). Repeated exposures to a novel word are the main vehicle of the incremental process of vocabulary acquisition, which begins with a lack of word knowledge, and proceeds to partial and finally full word knowledge (Frishkoff, Perfetti & Collins- Thompson, 2011; Joseph, Wonnacott, Forbes & Nation, 2014; Schmitt, 2010). Word knowledge is a multidimensional construct involving – according to Nation (2013) – nine components incorporating form, meaning (including the form-meaning link), and use (e.g., grammatical functions and contextual use). Fully mastering a word requires receptive and productive knowledge in each of these components. Generally, receptive word knowledge is formed earlier than productive knowledge (Pellicer-Sanchez, 2015; Webb, 2007). As well, learning gains may not necessarily be the same among the different components of word knowledge. For example, Godfroid et al. (2017) found that exposure has a stronger effect on the acquisition of word-form knowledge than it has on the acquisition of the form-meaning link. This complex structure of word learning is one of the reasons why it is necessary to determine the amount of exposures needed for mastering all aspects of word form and meaning. Although this question has been studied extensively, relevant studies report mixed results (see above). At least partly this is because the impact of exposure on learning quality is contingent on the type of reading materials, one's proficiency in reading the language (especially for L2 learners), the length, diversity and informativity of texts, and multiple additional factors (Elgort, Brysbaert, Stevens, & Assche 2018). Studies reviewed below have differentiated between *sufficient* and *maximum* number of exposures needed to facilitate vocabulary learning within the number of exposures they tested. *Sufficient* vocabulary learning corresponds to a minimal number of exposures needed for above-chance word knowledge, while *maximum* vocabulary learning is a number of exposures after which point learning does not further improve in the framework of the study.

One of the earliest studies using naturalistic reading instructed participants to read a novel in their L1 containing 241 nonwords (Saragi et al., 1978) and suggested that ten exposures are needed to sufficiently retain the meaning of a novel word. Similarly, Godfroid et al. (2017) used a novel containing 29 Dari words occurring 1 – 23 times and found that around 8 – 10 exposures are sufficient to enable meaning acquisition; yet, semantic and orthographic word knowledge can improve even after 23 exposures.

A book-reading study by Pellicer-Sánchez and Schmitt (2010) further showed that a substantial gain in form recognition, meaning recognition, and meaning recall tests at 5 – 8 exposures, while 28 or more exposures led to a gradual increase in all aspects of word learning. Mohamed's (2018) study using a graded reader reports that after 11 – 12 exposures, the novel words were read similar to the familiar words, and by 30th exposure the reading rate levelled off and did not increase further. Furthermore, Teng (2016) tested English as a foreign language (EFL) students in China using a graded reader which had nonwords occurring 1 – 20 times. They found that orthographic forms were learned sufficiently after 8 – 10 exposures, whereas semantics was learned sufficiently after 14 – 16 exposures. Participants reached their optimal performance after 18 – 20 exposures in each vocabulary post-test.

An oft-used alternative to natural reading is a word learning paradigm in which researchers directly manipulate the number and the context of the novel words. Creating the text stimuli allows the researcher to have more control over the context that the word occurs in, so factors such as readability and informativity of context, as well as frequency of occurrence of known words and the number of exposures to words can be accounted for or factorially manipulated. For instance, Hulme, Barsky, and Rodd (2019) created four short stories that embedded the target words two, four, six, or eight times. They found that the meanings of the words were recalled sufficiently after only 2 exposures, but accuracy became increasingly better with more exposures and peaked after 8 exposures. To eliminate the impact of context on orthographic learning, Ginestet, Valdois, Diard, and Bosse (2020) presented novel words in isolation either one, three, or five times. Both online eye movement monitoring and the offline orthographic choice test results suggest that orthographic learning begins very early, from the first few exposures of a word: two or three exposures are sufficient for orthographic learning. Yet, learners benefitted from extra repetitions as novel words presented five times were read and recognized faster. All in all, findings above generally converge on the estimate given in Nation's (2001) review that a range of 5 – 20 exposures are needed to learn a word (but see discussion in Elgort et al., 2018).

While all studies have reported that more exposures lead to better word learning, they diverged in their estimates of the number of exposures that is *sufficient* for gaining a non-trivial degree of orthographic or semantic word knowledge (defined as performing in a test above chance). Even fewer have assessed how many exposures to a novel word are needed for gaining a stable mental representation for the word's spelling or meaning, which does not substantially improve upon additional exposures, i.e., the maximum performance. Yet identifying a level of exposure to a novel word that affords maximum performance is important because it may shed light on the mechanisms responsible for the efficiency of human learning. The present study pursues this question.

Selective Attention

Noticing and memory encoding of the word must occur for word learning to emerge (Schmidt, 2001). Noticing is defined as the "conscious registration of the occurrence of some event" (Schmidt, 1990) and is subserved by the cognitive facility of selective attention. Bradley's (2009) review distinguishes three functions of selective attention: (i) detection of novelty and significance of the stimulus, (ii) enhancement of perceptual processing, and (iii) preparation for action (see, among many others, reviews by Pashler, Johnston, & Ruthruff, 2001; Posner and Petersen, 1990; Requin, Brener, & Ring, 1991; Seligman, 1970). In the case of novel word learning, these functions have been studied in related but separate lines of experimental research.

Selective attention to novel words has often been manipulated with a help of an orienting task, i.e., a task that draws a reader's attention to either the form or the meaning of a word in a text. Our discussion below draws from the review by Hulstijn (2003). For instance, a non-semantic orienting task may instruct readers to remember capitalized words in a text: in post-tests, participants would show better performance both on capitalized words (intentional learning) and other words with different visual features that they were not instructed to attend to, such as bold or italics (incidental learning). An example of a semantic orienting task would be to rate target words on their pleasantness: post-tests revealed a superior word knowledge for the rated items than non-rated ones. The nature of the orienting task (e.g., formal vs semantic) further predicts a greater success in post-tests that tapped into the knowledge of the novel word's form or meaning, respectively (Bransford, Franks, Morris, & Stein, 1979). Typically, semantic orienting leads to stronger learning gains than the non-semantic one (Hyde & Jenkins, 1973). This family of experiments can be argued to manipulate familiarity with different aspects of the stimulus and thus modulate its novelty and significance for the reader, in line with function (i) of selective attention.

Another family of experiments manipulates selective attention by implementing enhancement of perceptual processing, see function (ii) above. A common example is the use of visual and semantic enhancement techniques (Baleghizadeh, Yazdanjoo, & Fallahpour, 2018; Song, 2007; Izumi, 2002; Watanabe, 1997). In visual enhancement techniques, the appearance of the target form is manipulated using text formatting (e.g., bolding, underlining, highlighting) to increase its perceptual salience in the input (see Izumi, 2002). Arguably, the perceptually salient area attracts more attention and leads to better learning (Song, 2007). Text can also be enhanced semantically by providing explicit additional semantic clues to words or phrases. Providing the glosses of the target words, for example, would facilitate a reader to notice the target form, and more so than when word definitions were given inside a sentence immediately following the target word (Watanabe, 1997). Existing studies show mixed results about what amount and type of selective attention is necessary to affect perceptual processing and, subsequently, word learning (see e.g., Baleghizadeh et al., 2018; Izumi, 2002; Kang, 2003; Rott, 2007; Yang, 2004; Yeo, 2002).

To our knowledge, the only relevant systematic experimental manipulation that tapped into function (iii) of selective attention, i.e., preparation for action, is a comparison of incidental vs intentional learning on the same stimuli. Under most definitions (see Hulstijn, 2003), this distinction boils down to whether participants are made aware that they will be tested on their word knowledge after the learning phase is complete. Generally, greater gains are expected in intentional rather than incidental learning tasks because awareness of upcoming tests arguably increases learners' motivation and enables them to direct selective attention to the items to be tested on (see review by Hulstijn, 2003). The present study proposes a more fine-grained set of experimental conditions that modulate the reader's awareness and preparedness to the task at hand. Specifically, we provide half of the participants with instructions about how many times a novel word was embedded in the upcoming paragraph. This manipulation does not change their familiarity with stimuli (function i) nor does it enhance the presentation of the target novel words (function ii). However, it can be argued to allow a participant to mobilize attentional resources for the upcoming trial and plan energy expenditure proportionally to how many learning opportunities that trial affords (Requin et al. 1991). We provide further details on this attentional manipulation and formulate our predictions below.

Our study utilized eye-tracking, a popular experimental paradigm for studying attention during reading (Rayner, 2009; Lowell & Morris, 2017). It records where the reader fixates in the read text and for how long. The "eye-mind link" hypothesis (Carpenter & Just, 1983) suggests that eye movements reflect cognitive processing during reading, and thus longer eye fixations on a word reveal a greater amount of selective attention allocated to that word. This explains a frequent combined use of eye-tracking and post-tests of word knowledge in studies of language learning (e.g., Godfroid et al., 2017; Elgort et al., 2018; Mohamed, 2018; Pellicer-Sánchez, 2016). While eye-tracking points to the real-time progression of word learning (e.g., Ginestet, Valdois, Diard, & Bosse, 2020), the post-tests quantify the outcomes of learning. A common finding in eye-tracking studies of word learning is that reading times are particularly long during the first encounter with the word but then decrease and finally reach an asymptote in further encounters (see above and a review by Winke, Godfroid, & Gass, 2013). This decrease occurs as early as after the first encounter in the L1 group whereas L2 readers require around three or four encounters before reaching a plateaued reading speed (Pellicer-Sánchez, 2016; Godfroid et al., 2017).

In sum, vocabulary learning has been argued to benefit both when the novel word is presented to the reader more times and when attention is allocated at each presentation (Nation 2015; Schmidt 2010). Yet it is presently unknown whether the effects of these two dimensions of interest – exposure and attention – are additive or interactive. Furthermore, the exact parameters of these effects are not fully explicated either for L1 or L2 learners. For instance, previous studies have not reached a consensus on whether there is an optimal number of exposures for acquiring orthographic or semantic knowledge in L1 or L2 novel word learning, and if so what this number is (see discussion above and in Hulme et al. 2019; Pellicer-Sánchez 2016).

The present study

We aim to contribute to the literature on novel word learning by implementing a factorial design which manipulates both attention and exposure. Our first goal was to determine what level of exposure leads to a sufficient orthographic and semantic knowledge (defined as performing above chance on respective tests) or to a maximum knowledge (defined as performance that does not improve after additional exposures) within the study's parameters. Our second goal was to identify whether selective attention plays a unique role in novel word learning and whether it modulates the impact of exposure. We pursued these goals in a study administered to L1 and L2 speakers of English in a novel-word learning paradigm. Participants read short passages with embedded novel words while having their eye movements recorded. The passages provided informative context for inferring the meaning of these novel words. The readers were then asked to complete surprise post-tests measuring semantic and orthographic knowledge of the novel words they read.

The combination of eye-tracking and post-tests enabled us to tap both into the real-time process of novel word learning and its orthographic and semantic outcomes. Additionally, we tested the immediate and delayed impact of our experimental manipulations on acquisition and retention of new vocabulary knowledge by conducting the vocabulary post-tests (defined below) both immediately after the learning phase and one week later.

The experiment had a 3 x 2 factorial design. Our first factor of interest, Exposure, was manipulated by a different number of occurrences of novel words, appearing 2, 4, or 8 times in a 9-sentence paragraph. A manipulation of attention can be induced in a variety of ways, reviewed above. We opted for inducing selective attention to occurrences of novel words by presenting instructions concerning the number of exposures to some participants, while letting other participants read texts with novel words naturally. Half of the participants both in the L1 and the L2 groups received attention-inducing instructions: "In the following passage, you will see a novel word 2 times/ 4 times/ 8 times", and the other half did not. The presence of instructions did not provide additional information about the form or the meaning of the novel word but could lead to more efficient mobilization of attentional resources in accordance to how many occurrences of a novel word will be afforded in the text (see above and Bradley, 2003; Requin et al., 1991). That is, one of the conditions contributed to a core aspect of selective attention, i.e., the reader's preparation for the action, by informing the reader about the expected complexity of the task. Figuratively speaking, our manipulation is similar to informing some (but not all) of participants in a hurdling race about how many obstacles they will have to jump over. We predicted that knowing that a word will only appear in a passage two, four or eight times, may enable the reader to make a strategic use of the smaller or larger number of learning occasions. The condition without instructions did not afford this possibility. If this manipulation of function (iii) of selective attention is effective, we expected participants to show greater learning gains when they receive instructions about the relative complexity of the upcoming task.

An argument can be made that awareness of the presence and number of novel words in one of the conditions alters the type of learning from incidental to intentional in this condition. Since neither condition reveals that word knowledge will be tested after the learning phase, most existing definitions would attribute both conditions to incidental learning (e.g., Hulstijn, 2003). Because the interpretation of our manipulation and its effect on learning are independent of specific definitions, we are agnostic to a possible terminological ambiguity.

In sum, we aim to examine effects of exposure and attention on the eye-movements registered during novel word learning and the results of immediate and delayed vocabulary post-tests in L1 and L2 learners of English. Our first goal is to determine the efficiency of our manipulation of selective attention to novel words and their implications for learning novel form and meaning. Our second goal is to determine, within an experiment, the number of exposures to a novel word required for a sufficient and maximum performance in post-tests. Finally, we examine whether these two major predictors of novel word learning – attention and exposure – interact either during the real-time learning phase or in the post-tests of orthographic and semantic knowledge.

Methods

Participants

Twenty-nine L1 undergraduate students and 29 L2 undergraduate students participated in this experiment. We removed six participants from the L1 sample and three participants from the L2 sample due to excessive blinking or the unsuccessful calibration of the eye tracker. The final sample consisted of 23 L1 (age M = 20.9, SD = 2.2; age range 17 to 28; 21 females) and 26 L2 (age M = 19.6, SD = 2.6; age range 17 to 28; 19 females) undergraduate students of a university in Canada. All participants were recruited through an online subject pool and rewarded by course credit. All participants in the L2 group declared English as their second language and their ART and Self Rating Scores were significantly lower than the L1 group (see Supplementary Materials S1 for details).

L2 participants represented 9 different language backgrounds: 17 were speakers of Chinese (including Mandarin and Cantonese), 2 spoke Gujarati, and seven additional languages were represented by 1 speaker each. The majority of L2 participants were Mandarin speakers enrolled in an EFL program for international students.

All participants had normal or corrected-to-normal vision and did not report any learning or visual impairments. This study was approved by the McMaster Research Ethics Board 2011-165.

Materials

The target stimuli for the reading task consisted of nine novel words (e.g., *plurk*). The novel words, generated using the Wuggy software (Keuleers &

Brysbaert, 2010), were chosen such that they had a homophone (e.g., *plirk*), see justification below.

To accommodate both L1 and L2 readers, the passages were designed to be syntactically simple and contain only frequent non-target words. Also, we only selected content words that occurred at least 1,000 times in the 100-million token BNC corpus. The texts were constructed in the genre of narrative prose. Each paragraph consisted of nine sentences. Our critical manipulation of Exposure to the novel word was implemented by embedding the target word in either 2, 4, or 8 initial sentences in each passage: a novel word would occur exactly one time in each of these sentences. Each passage concluded with a semantically related sentence that did not contain the target word (for stimuli, see Supplementary Materials S3), for example:

To spend an afternoon outside, Nicole used a plurk to dig small holes in the garden. The dark brown handle of the plurk was made of a light wood. The smooth top part of the plurk was easy for her to hold. The sharp metal blade of the plurk helped Nicole complete her work really fast. She relied on the plurk to plant this space with different vegetable seeds. She admired how convenient the plurk was for her work. After she was done, she immediately put the plurk in the shed. This was because she did not want her plurk to get rusty in the backyard like last time. Nicole always takes care of her belongings to avoid unwanted damage.

The corresponding passages with 2 (4 or 8) exposures to a novel word would contain the same 2 (4 or 8) initial sentences as in the example above and thus would give 2 (4 or 8) exposures to the novel word, respectively. The same concluding sentence followed the experimental sentences. The remainder of the 9-sentence passage would consist of 6 (for the 2-exposures condition) or 4 sentences (for the 4-exposures condition), none of which would include a target novel word. Also, we made the non-critical sentences constituting the remainder of each passage semantically unrelated to the content of target sentences. This way we avoided rehearsing the form or strengthening the learning of the meaning of novel words. The number of occurrences of each novel word was counterbalanced to ensure that no participant saw the same word in two different Exposure conditions. All novel words had definitions and syntactic roles corresponding to common concrete nouns (e.g., tool, clothing, device). None of those words were placed at the end or beginning of a line.

Procedure and apparatus: For the eye-tracking phase, participants were seated in a comfortable chair approximately 65 cm in front of an NEC MultiSync LCD 17-inch computer monitor with a resolution of 1600 x 1200 and screen refresh rate of 60 Hz. Courier New 20 point fixed-size was used for text passages, resulting in about 3 characters subtending 1 degree of visual angle. Eye-movements were recorded with an EyeLink 1000 desktop eye-tracker (SR Research, Kanata, Ontario, Canada) at a sampling rate of 1000 Hz. A 9-point calibration was set to be performed at the beginning of the experiment. Stimuli were viewed binocularly, but eye-movement data from only one eye was analyzed. Drift correction took place at the beginning of each trial and calibration was monitored and redone by the researcher if necessary.

Participants were tested over two sessions one week apart. In the first session, participants filled out brief demographic and language background questionnaires and individual differences tests, see Supplementary Materials S1 for individual differences measures and Table S1 for self-rated proficiency. Then they read passages while having their eye movements tracked. Finally, they completed post-tests including orthographic choice, definition matching, and definition prompting tasks, in this order. In the second session, participants completed a second set of the same post-tests: no additional exposure to novel words was provided. The first session did not exceed one hour; the second session took half an hour. Below we describe each task in detail.

Passage reading: Each passage appeared on a separate screen and occupied nine lines. Participants were instructed to press a key when their reading of a passage was completed. The order of passage presentation was randomized for each participant. All participants were asked to read paragraphs carefully and silently for comprehension even as they encounter unfamiliar words. Participants were assigned to two lists to receive different instructions: half of the participants were informed about how many times a novel word would appear in the upcoming passage prior to reading each passage, while the other half were not.

Vocabulary post-tests: These tests measured the outcomes of learning the novel word's orthography and meaning, as inferred from the passage reading.

The *orthographic choice* task involved a computer mediated forced choice test: to test orthographic knowledge, participants were instructed to discriminate between the target novel words and filler homophones (e.g. novel word *plurk* filler *plirk*). Also, their reaction times in milliseconds were recorded during this task.

In the *definition matching* task, the participant received a list of the nine target novel words, along with nine unseen foil words. Similarly, there were nine correct definitions and nine foil definitions. Participants were instructed to match the target novel words that they read to the correct definition. In this task, accuracy indicated association of newly-learned forms with meanings inferred from context.

The *definition prompting* task is a measure of productive semantic knowledge. It was administered verbally, where participants were asked to produce the meaning of a word upon hearing it. Nine unseen foil words were also included. The questions were asked by the experimenter in a random order one at a time. The participant was aware of the presence of the foil words and they were instructed to skip them by stating "That was not one of the words that I read". Participants were allowed to ask for one hint for the seen novel words: hints were segments of the passage that they read containing this word. Only correct responses to seen novel words were scored, 2 points for the responses on the first attempt and 1 point for the responses after a hint.

Variables

Reading: We selected total fixation time on the word (a summed duration of all fixations on the word) as a cumulative measure of cognitive effort of word recognition which has been also suggested as a strong predictor of word learning (Godfroid et al., 2017; Godfroid et al., 2013; Mohamed, 2017; Pellicher-Sanchez, 2015). We considered this measure both in aggregation for every level of Exposure (e.g., total fixation time on the novel word averaged over two, four, or eight word occurrences) and also for every specific occurrence (e.g., total fixation time to the novel word at its first, second, third, and fourth occurrence in a 4-occurrence condition of Exposure). This enabled us to tap into both the overall impact of our experimental conditions on eye-movement patterns and a detailed report of how these conditions influence every specific learning opportunity.

As a control variable, we also considered the time spent reading each passage. The total passage reading times indicated the amount of attention paid to the entire passage. Since introducing (logged) total reading time for novel words did not have any significant effect on any of the post-test performances, we do not include it to the reported models below (see Supplementary Materials S2, Table S2 for descriptive statistics).

Word knowledge: Additional dependent variables represent accuracy scores from all vocabulary post-tests and RTs from orthographic choice task (defined above) obtained both during the first immediate testing session and the second session delayed by one week. Only responses to the target novel words were included in the analyses.

One of the critical manipulations of this study is the reader's awareness of the presence and number of novel word's occurrences in the upcoming passage. We expected awareness and preparedness to the task to influence both the eye-movements to the novel word and learning outcomes of readers on vocabulary post-tests. This manipulation, Instruction, was treated as a categorical variable with two levels (presence vs absence of attention-inducing instructions). The number of exposures to a novel word was the other experimental manipulation of this study. A novel word appeared either 2, 4, or 8 times in a passage. Exposure were treated as a categorical variable with three levels. Finally, Session was introduced as a categorical variable with two levels (immediately after learning or delayed by a week, also see Statistical Considerations for the details).

Statistical considerations We fitted linear mixed effect regression models to the eye-movement measures and generalized linear mixed effect regression models (using the lme4 package, Bates, Maechler, Bolker, & Walker, 2015, in the statistical software environment R, R Core Team, 2018) to the rest of dependent variables described above using as independent variables Session, Instruction, and Exposure, as well as by-subject and by-word random intercepts. The lmerTest package (Kuznetsova et al., 2017) was used to estimate p-values for fixed-effects with Satterthwaite's approximation for degrees of freedom. Post-hoc comparisons were performed using the emmeans package (Lenth et al., 2019). The Gaussian distribution family was used in models fitted to continuous variables, the Poisson family to count variables (e.g., test scores ranging from 0 to 2), and the binomial family to binary variables. As recently recommended by Lo and Andrews (2015), the Gamma family and the identity link were used for reaction times (RTs) data to account the right-skewed shape of the distribution without any need for data transformation.

A total of 12 models were fitted, 6 for the L1 and 6 for the L2 cohort. In each group of 6 models, three tapped into accuracy results in each of the post-tests (see above). An additional model analyzed RTs of the orthographic choice task. Two more models were created to analyze eye tracking data: One for eye-movements aggregated over all novel word occurrences in the levels of exposure condition and one for eye-movements registered during the individual occurrences of novel words. The critical effects on all independent variables were those of Exposure to the word (manipulated as 2, 4, or 8 word occurrences in the passage) and Instruction (manipulated via presence or absence of attention-inducing instructions), as well interactions between the two factors. Since vocabulary learning measures were administered twice, one-week apart, Session was also introduced to those models as an independent variable. After fitting eye-tracking models, we removed outliers with standardized residuals outside the interval [−2.5, 2.5] and refitted the model.

Log-likelihood ratio model comparisons were conducted using the anova function in R. The initial models included the main effects and the random effects. Since we were interested in whether the effects of attention and exposure are additive or interactive, the interaction terms were introduced in turn between two main effects keeping the random-effects structure intact. The interaction model was selected as optimal when its performance was significantly higher than that of the model without the interaction term.

As a statistically insignificant outcome may not be informative regarding the absence of an effect (e.g., the amount of support for rejecting or failing to reject the null hypothesis), we also conducted Bayesian analyses in cases where null effects were theoretically important. The Bayes Factor quantifies evidence both for the null hypothesis (with no effect), and the alternative hypothesis (with desired effect). We computed the Bayes Factor by calculating the difference between the Bayesian Information Criterion (BIC) values of these two models, using the following formula as proposed by Masson (2011):

$$BF = e^{(BIC_0 - BIC_1)/2}$$

BIC_0 represents goodness-of-fit of the null model and BIC_1 that of the model including the effect of interest. The BIC values for each regression model were obtained by using the BIC function in the base distribution of R. A Bayes Factor greater than 3 represents moderate and greater than 6 represents strong evidence for the alternative hypotheses whereas values below 1/3 moderate and below 1/6 show strong support for the null hypothesis. Finally, a Bayes Factor between 1/3 and 3 gives indecisive evidence for or against a hypothesis.

Results and Discussion

To anticipate the presentation of results, the manipulation of Instruction had a crucial impact on processing speed. However, its influence on learning quality was limited: we offer possible explanations for these null results in the General Discussion. We present results of the experiment by task, emphasizing in each analysis the role of Exposure and Instruction as critical variables. We first summarize the eye-movement behaviour during reading passages with novel words, and then report the vocabulary post-tests for L1 and L2 speakers of English.

Eye movements

Twenty-three L1 participants contributed 25,779 data points, of which 966 were on target words. We further removed trials in which target words were skipped (100 data points) and elicited very short fixations (< 80 ms), very long fixations (> 1500 ms), or a very large number of fixations (>5). The resulting data pool consisted of 845 data points. Twenty-six L2 participants contributed 29,143 data points, of which 1,092 were to the target words. We further removed trials in which the target word was skipped, or the fixations on the target word were too short or too long (see above). This trimming left a pool of 800 data points.

The first step of analysis considered total fixation times to novel words aggregated over conditions of Exposure: i.e., total fixation times registered in the 2-exposures condition were considered jointly, regardless of whether they came from the first or the second occurrence of the novel word. Figure 1 visualizes total fixation times broken down by conditions of Exposure and Instruction, for L1 (left panel) and L2 readers (right panel). Among L1 readers, total fixation times were shorter in the instruction rather than no-instruction condition: This suggests that the instruction manipulation did induce selective attention to novel words. Total fixation times also decreased with a greater overall number of exposures: e.g., both in the instruction and no-instruction conditions, novel words that occurred twice were looked at longer on average than the words occurring four or eight times in the passage. This decrease related to Exposure was more salient in the no-instruction conditions than in the passages where instructions potentially induced selective attention. In conditions with 2 or 4 repetitions of a novel word, the presence of attention-inducing instructions led to faster reading of those words: we did not find the same advantage in the condition with 8 exposures. Overall, this finding emphasizes the role of attention when the number of learning opportunities with the novel word is limited. A regression model fitted to total fixation time (summary reported in Supplementary Materials S2, Table S3) in L1 speakers confirmed an interaction between Exposure and Instruction.

In L2 readers, average total fixation times (ranging roughly between 800 in the 2-exposures condition to 500 in the 8-exposures condition) were overall longer than those recorded in L1 readers (ranging from roughly 500 to 320 ms, respectively). Similar to L1 readers (Figure 1 left), a larger number of exposures to the word led L2 readers to a consistent speed-up. On average, L2 readers were faster to read a word that occurred 8 times versus 4 times or 2 times. In contrast to L1 readers, the Instruction manipulation did not affect total fixation times, nor was there an interaction between Instruction and Exposure, see Supplementary Materials S2, Table S4 for an output of the regression model.

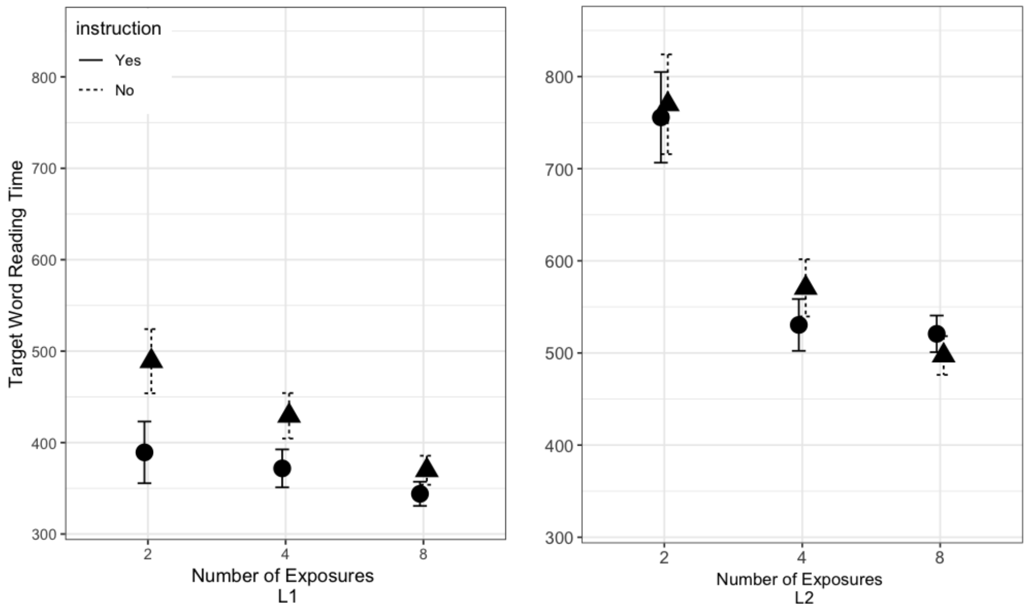

Figure 1. L1 (left) and L2 (right) total fixation time by conditions of Instruction and Exposure

To zoom in on the fine-grained behavioral patterns across multiple exposures to the novel word, our follow-up analysis considered eye-movements registered during the individual occurrences of novel words, rather than the eye-movements aggregated over all novel word occurrences in the 2-, 4- or 8-exposures conditions. Figure 2 (left) reveals that in L1 readers mean total fixation times to the first occurrence of a novel word were significantly longer than total fixation times to each of the other exposures. Exposures after the first one did not differ reliably from one another in the 4- and 8-exposures conditions (model not shown). Thus, in L1 speakers the crucial impact on processing speed is obtained from the first exposure to the novel word while all additional exposures to the word only lead to a largely asymptotic high processing speed. These findings converge with a recent report by Ginestet et al. (2020) which also indicated an early sharp decrease in reading times in an experiment using a different presentation of novel words. Figure 2 (left) further highlights the impact of Instruction: this impact concentrates on the early occurrences of the novel word. It affects both occurrences of the novel word in the 2-exposures condition of Exposure where the presence of instructions comes with shorter total fixation times. The instruction condition speeded up eye-movements to the first occurrence of the novel word in the 4- and 8-exposures condition as well: There was no consistent contrast in average total fixation times to other novel word occurrences in these conditions (model not shown).

In L2 readers, each occurrence in each Exposure condition led to a decrease in total reading times on the novel word. This suggests a similar recognition and learning mechanism of novel words across L1 and L2 readers and across different types of stimuli (the one used in this study and that in Ginestet et al., 2020). The Instruction manipulation did not significantly affect total reading times at any occurrence of the novel word (model not shown).

To summarize the eye-movement analyses, manipulation of Exposure led to consistent and expected overall trends in the processing of novel words. Specifically, the more occurrences of a novel word both L1 and L2 participants had a chance to experience, the faster the overall reading times were on that word. Yet the speed advantage coming with multiple exposures to a word was mostly gained during the first exposure in L1 readers, while L2 readers benefitted from additional exposures up to the first four occurrences of the novel word. The presence of instructions in one of the Instruction conditions did induce selective attention to novel words, which showed in shorter average reading times to the novel words when their expected number was made known to the reader. However, this effect was only observed in L1 readers and only in their first exposure to the novel word. The L2 readers were not affected by our manipulation of selective attention.

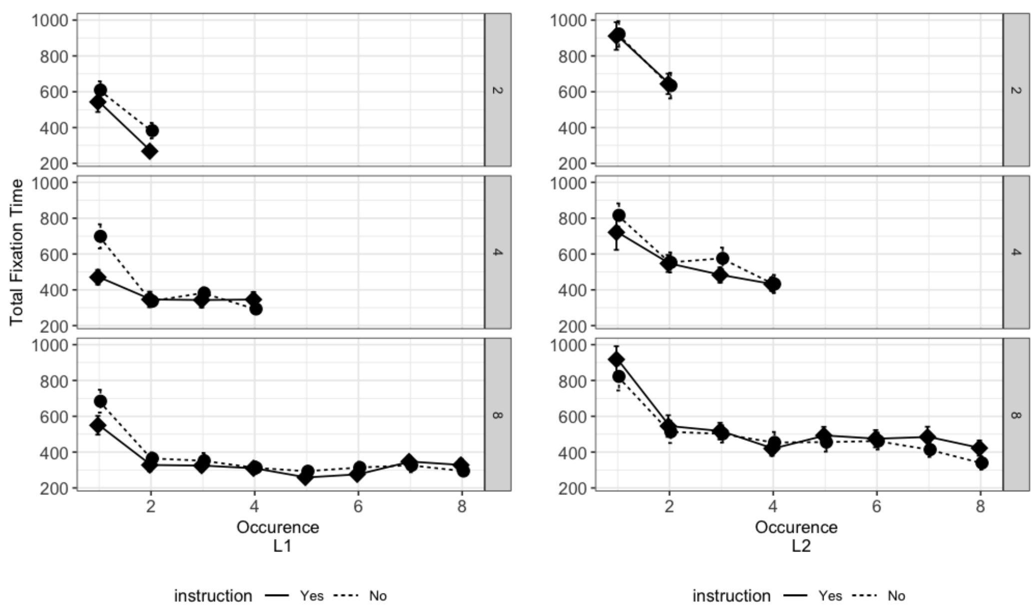

Figure 2.Total fixation times to every occurrence of the target in L1 (left) and L2 (right) data, by conditions of Exposure and Instruction

Vocabulary Post-tests The tests below assess the outcomes of learning the novel word’s form and meaning.

Orthographic choice task

In this analysis we consider responses to the seen novel words only. For L1 speakers, a total of 450 responses were recorded; and a total of 558 responses were recorded for the L2 speakers. Figure 3 displays average accuracy of orthographic choice responses by testing session (immediately after learning or delayed by a week) and Exposure (a novel word occurring a total of 2, 4, or 8 times in the passage). Figure 3 (left) shows that testing session had a strong main effect on orthographic accuracy in L1 speakers, with the delayed responses in session 2 being consistently more accurate than in session 1. This main effect is confirmed in the mixed-effects logistic regression model reported in Supplementary Materials S2, Table S5. Exposure was another reliable predictor of orthographic accuracy. The lowest accuracy was observed in the 2-exposures condition. As indicated by the post-hoc proportion test, responses in session 1 were not significantly different from the baseline of 50% correct (observed 47%, $\chi^2 = 0.12$, df = 1, $p = 0.734$). Responses in session 2 of this condition (68% correct) and in all other conditions were reliably more accurate than the baseline. Importantly, the maximum accuracy for respective sessions was achieved in the 4-exposures condition (session 1: 65% and session 2: 79% correct). Additional exposures in the 8-exposures condition did not lead to a statistically significant increase in accuracy (session 1: 69% and session 2: 78% correct; all *p*s > 0.05). Thus, exposure to as few as two occurrences of a novel word in a passage translate into an above-chance performance in a session administered one week after learning in L1 readers. Furthermore, exposure to four novel word occurrences was optimal for L1 readers to encode and retain the orthographic form of the novel word given the parameters of the present experimental study.

Figure 3 (right) summarizes average orthographic choice accuracy by Session and Exposure in the L2 cohort. Unlike the L1 sample, the 2-exposures condition did not lead to an above-chance performance in this task (session 1: 50% and session 2: 58%; both *p*s of the proportion test with the 50% baseline > 0.1). All other conditions came with scores reliably above the baseline (all *p*s in post-hoc proportion tests < 0.01). An increase in the number of exposures led to a consistent increase in accuracy: the 8-exposures condition came with higher accuracy (session 1: 81 % and session 2: 80%) than the 4-exposures condition (session 1: 70% and session 2: 73%), and accuracy in both was higher than in the 2-exposures condition. This suggests that for L2 readers all additional exposures were beneficial for orthographic learning. Testing session did not affect L2 readers. All trends reported above were confirmed by the mixed-effects logistic regression model fitted to responses in the orthographic choice task, see Supplementary Materials S2, Table S6.

Within our testing paradigm, four exposures to the novel word were the minimum exposure necessary to encode the novel word's orthography, while the maximum performance for this experiment was shown in the 8-exposures condition. It is noteworthy that orthographic accuracy in the L2 sample after either 4 or 8 exposures was comparable or numerically higher than that demonstrated by the L1 sample.

The Instruction manipulation did not have any main effect on the accuracy of orthographic choice ($BF_{L1} = 0.05$, $BF_{L2} = 0.04$), nor did it interact significantly with Session ($BF_{L1} = 0.11$, $BF_{L2} = 0.06$). Both groups did not show any interaction effect between Instruction and Exposure ($BF_{L1} = 1.29$, $BF_{L2} = 0.01$). The Bayes factor between 1/3 and 3 suggests that the evidence is inconclusive to support or reject that the interaction between Instruction and Exposure has an effect on orthographic accuracy for the L1 sample. All other Bayes Factors less than 1/6 suggests the presence of strong evidence for the null over the alternative hypothesis.

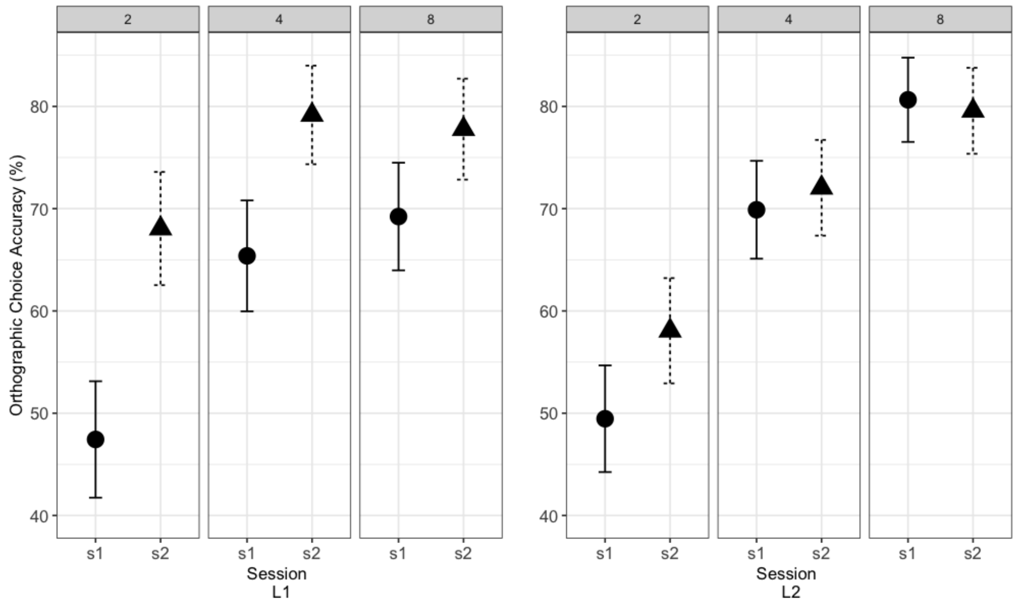

Figure 3. L1(left) and L2 (right) Orthographic choice accuracy scores per Session broken down per level of Exposure

We further analyzed response times of the orthographic choice decisions to seen and correctly identified novel words, see Figure 4. Figure 4 (left) displays that the L1 sample responded faster when they were tested immediately after the reading task than when they were tested one week later (session 1: M = 1194, SD = 754 and session 2: M = 1336, SD = 866). Figure 4 also displays an interaction between Session and Instruction. Those who received the instruction responded faster in the first session (the instruction condition: M = 1052, SD = 402 and the no-instruction condition: M = 1336, SD = 969), however, this advantage disappeared when they did the same test one week later (the instruction condition: M = 1421, SD = 867 and the no-instruction condition: M = 1251, SD = 860). These effects are confirmed in the mixed-effects regression model reported in Supplementary Materials S2, Table S7. We conducted an additional contrast to examine the effect of Instruction separately for each session. The results also confirmed that the group receiving the instruction responded faster in the first session (p = 0.004), this effect is reversed in the second session with faster responses from those who did not receive the instruction (p = 0.03), see Figure 4 left and Supplementary Materials S2, Table S8.

Figure 4 (right) displays estimated values of RTs in the orthographic choice task per levels of Session and Instruction for the L2 sample. Like the L1 group, RTs were faster in the first session (M = 1360, SD = 1142) compared to the second session (M = 1453, SD = 1269). However, different from the L1 cohort, the presence of attention-inducing instruction during the reading task led to a persistent advantage in both sessions with shorter RTs among the L2 group (the instruction condition: M = 1274, SD = 902 and the no-instruction condition: M = 1530, SD = 1426). The exposure manipulation also had a significant effect on RTs of orthographic choice task: the L2 readers responded faster to words presented in the 4-exposures condition (M = 1278, SD = 642) than those presented in the 2- (M =1465, SD = 1281) and 8-exposures (M = 1476, p = 1519) conditions. All reported trends were confirmed by the mixed-effects logistic regression model, see Supplementary Materials S2, Table S9.

Although it did not have any effect on orthographic accuracy results, the Instruction manipulation influenced orthographic choice RTs for both groups: This effect led to short-lived speed up in RTs in the L1 sample whereas it provided a reliable benefit to the L2 sample with shorter response times in both sessions.

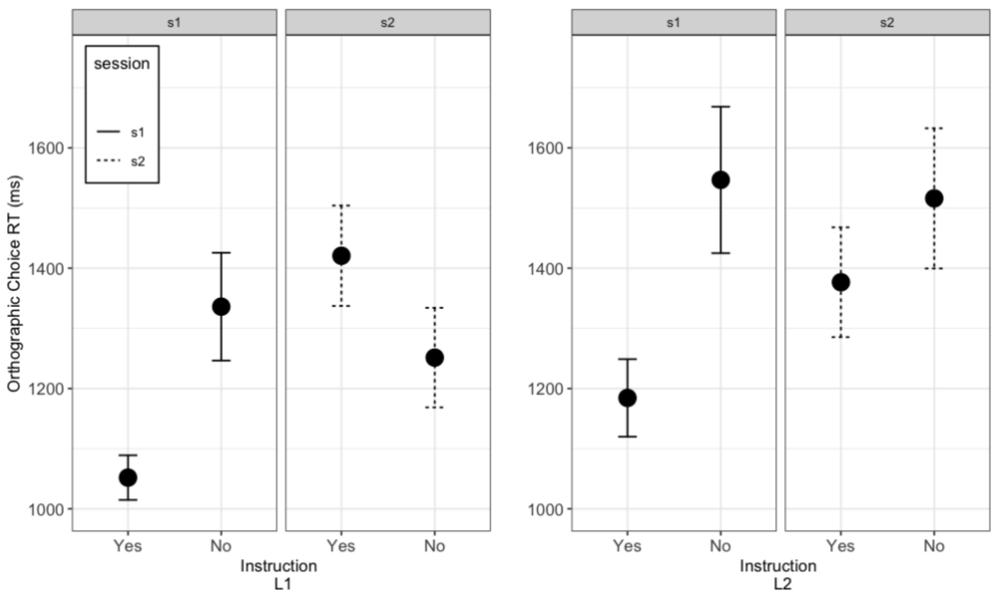

Figure 4. L1(left) and L2 (right) Estimated values of RT per Session and Instruction

Definition matching task

A total of 413 responses to the seen novel words were recorded in the L1 sample, and a total of 459 responses in the L2 sample. Figure 5 summarizes average definition matching scores per Instruction broken down per level of Exposure: since the manipulation of session did not affect these outcomes, it is not visualized for either sample. Figure 5 (left) suggests that the accuracy of the definition matching task in the L1 sample increased from the 2-exposures condition (14%) to the 4-exposures condition (22%) and did not further increase in the 8-exposures condition (21%). The baseline performance in this test, achieved if one were to match the seen items and definitions randomly, is 5.6%. Accuracy shown in the 2-exposures condition was significantly above the baseline (14%, $\chi2 = 16.036$, df = 1, $p < 0.001$). Thus, for L1 speakers of English two exposures were sufficient to gain a degree of semantic knowledge about the novel words they saw. After 4 exposures, they reached their optimal semantic knowledge, which did not improve with additional exposures to the novel word within the experiment (22% and 21% respectively). The mixed-effects logistic regression model reported in Supplementary Materials S2, Table S10, as well as post-hoc analyses of its coefficients, supported the patterns above. The model also confirmed that session or presence of instruction did not have a reliable effect on definition matching scores. The computed Bayes Factors less than 1/6 also provided strong evidence in favor of the null hypotheses: Neither the presence of instruction (BF = 0.05), nor its interaction with Session (BF = 0.05) or Exposure (BF = 0.002) did not play any role in semantic recognition performance of the L1 group.

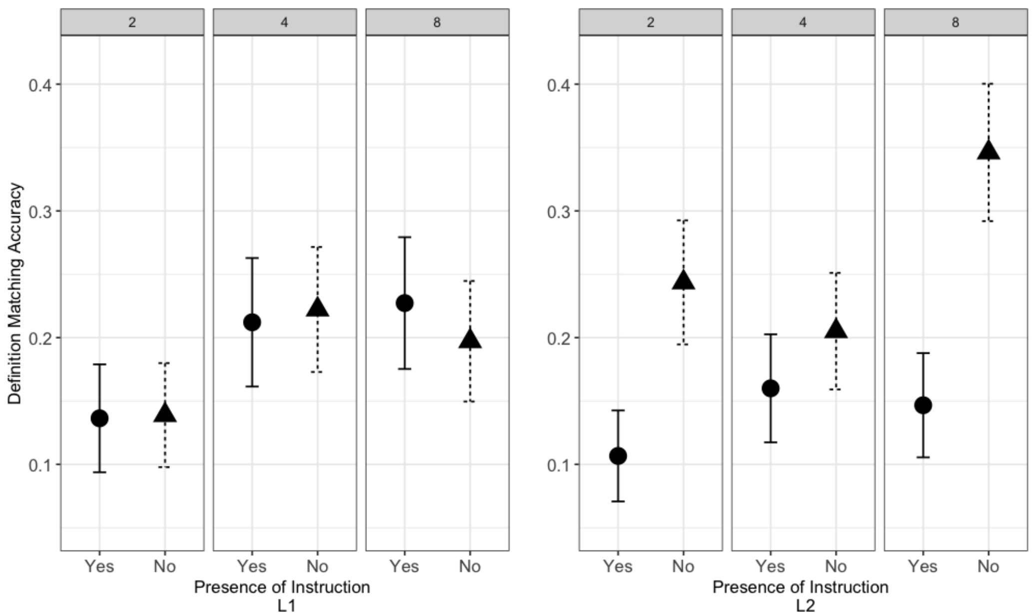

Figure 5. L1 (left) and L2 (right) Definition matching scores per Instruction broken down per level of Exposure

Figure 5 (right) displays the average definition matching scores across levels of Exposure for L2 speakers. They showed a similar performance in this task of semantic knowledge after either 2 or 4 exposures (18% and 18%) but improved their performance after additional exposures (8-exposures: 25%). It is noteworthy that even after 2-exposures, L2 participants showed an above-chance level of semantic knowledge (as indicated in the proportion test against the 5.6% baseline, $p < 0.01$). Their performance in this task was also numerically higher than that shown in the corresponding task by L1 speakers.

The mixed-effects logistic regression model fitted to the task responses (Supplementary Materials S2, Table S11) showed that the numeric advantage of the 8-exposures condition over other conditions is above the nominal threshold of significance ($p = 0.087$). It also points to a counterintuitive reliable effect of Instruction: when L2 readers do *not* receive attention-inducing instructions in the learning phase, they are more successful in acquiring semantic knowledge from the texts [b = 1.088, SE = 0.503, z = 2.163, $p = 0.031$]. L2 readers' average performance nearly doubled, increase from 13% to 26%, when their attentional resources were not occupied by the knowledge of how many times a novel word would occur in the text (see Figure 5). The week-long delay between sessions did not affect this task's performance. The Bayes Factor also provided substantial evidence in favor of the absence of interaction effect between Instruction and Exposure (BF = 0.08) and Instruction and Session (BF = 0.05).

Definition prompting task

For L1 speakers, a total of 414 responses to the seen novel words were collected, and a total of 459 responses were collected for L2 speakers Figure 6 (left for L1 and right for L2 speakers) shows the average definition prompting scores per conditions of Exposure. Among L1 speakers, the score was the lowest in the 2-exposures condition (0.28, 14%), the highest in the 4-exposures condition (0.40, 20%) and in the mid-range in the 8-exposures condition (0.33, 16%). The baseline for this task, which would be achieved if definitions to novel words were generated randomly, is a score of 0. Thus, even the lowest score in the 2-exposures is significantly higher than the baseline ($p < 0.05$). Considering the results of definition matching and the definition prompting tasks together, we conclude that 2-exposures to the novel word grant L1 participants with a non-trivial (above-chance) amount of semantic knowledge, while 4 exposures are optimal for semantic knowledge acquisition and is not improved on by additional exposures. The mixed-effects regression model fitted to the task scores (Supplementary Materials S2, Table S12) identified the contrast between the 2- and 8-exposures condition as highly reliable, while the contrasts between other pairs of condition were not significant. As with the matching task, this suggests that for L1 readers the 4-exposure condition is best for semantic learning. The model (see Supplementary Materials S2, Table S12) additionally revealed null effects of Instruction and Session on the scores. The Bayes factor also suggested strong evidence for the null effect of Instruction (BF = 0.06), and the absence of interactions between Instruction and Exposure (BF = 0.01), and Instruction and Session (0.05).

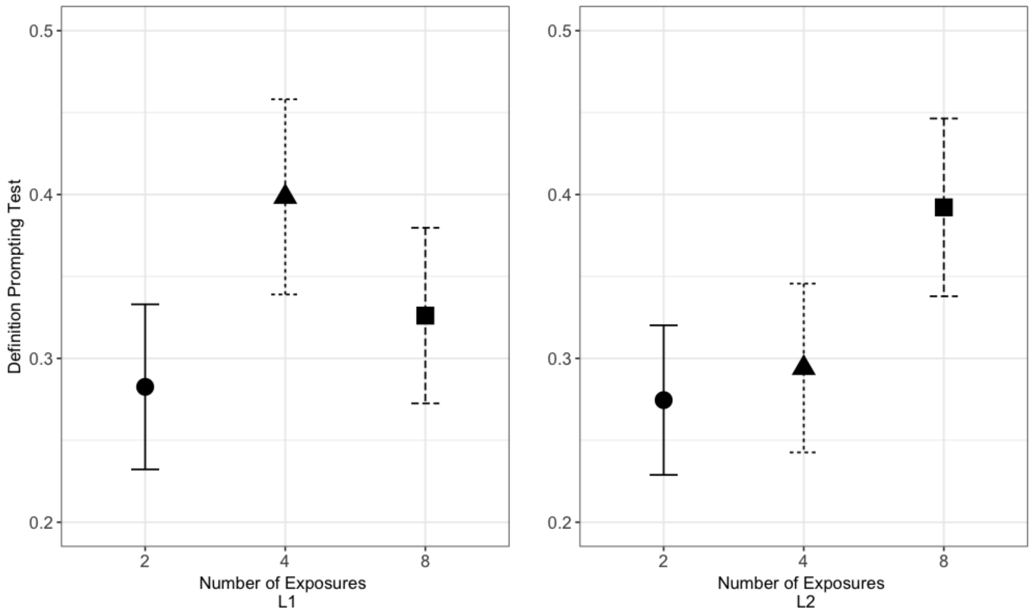

Figure 6. L1 (left) and L2 (right) Definition prompting scores per level of Exposure

In the L2 sample, average scores for the 2-, 4- and 8-exposures conditions were 0.27 (14%), 0.29 (15%), and 0.39 (20%), respectively, see Figure 6 (right). Similar to other tasks, the performance of L2 participants in this definition prompting task was comparable to that in the L1 sample. As with the orthographic choice and definition matching task in this cohort, the highest performance in this population was shown after the maximum number of exposures to the novel word. Also, similar to the other semantic knowledge task, two exposures were sufficient to engender an above-chance performance ($p < 0.05$). The regression model fitted to the definition prompting score showed that neither main effects, Exposure, Instruction and Session nor their interaction influenced the definition prompting score (see Supplementary Materials S2, Table 13). The computed Bayes Factors were all smaller than 1/6 and thus strongly supported a null effect of Instruction (BF = 0.05), Exposure (BF = 0.01), and their lack of interaction with each other (BF = 0.04) and with Session (BF = 0.06 and BF = 0.004, respectively).

General Discussion

The main focus of this paper is on two proposed causal factors of successful novel word learning deeply grounded in theories of word learning: the amount of experience with these words and the amount of selective attention to the words. Exposure only captures the number of learning opportunities with the novel word. An important additional dimension is how efficiently these learning opportunities are utilized; this, in turn, depends on the degree of selective attention to the novel word (e.g., Schmidt 1990; 2001). We were interested in what combinations of exposure and attention would initiate word learning (i.e., above-chance performance in post-tests) and lead to the maximum performance within the number of exposures tested in this study (i.e., performance that does not improve through additional exposure or induced attention).

The present study addressed these questions by presenting L1 and L2 speakers of English with passages to read in a novel word learning paradigm. Exposure was manipulated as the number of occurrences for a novel word in the passage (2, 4, or 8). To control the degree of attention to novel words, we presented participants with instructions that either specified the number of novel word occurrences in the upcoming passage or did not. Eye-movements of readers were monitored during the learning phase, while surprise tests of orthographic and semantic knowledge of novel words measured learning outcomes immediately and a week after.

We see the main contribution of this paper in that it considers the two factors (Exposure and Instruction) simultaneously in both L1 and L2 readers of English, using both online and offline indicators of word learning. This enabled us – within one study – to quantify and directly compare the effect of either factor on two populations of novel word learners, pin down where in the processing of learning these effects take place, and estimate whether these effects are additive or interactive. We discuss our findings below.

Exposure

To our knowledge, ours is one of the few studies that examined the effect of exposure frequency by presenting a varying number of occurrences of a novel word within a short story presented on a single screen, as opposed to spreading these occurrences over multiple segments of a larger text (Godfroid et al., 2017; Mohamed, 2018; Pellicer-Sánchez, 2016). This is similar to the Wikipedia type of reading when a single encyclopedic entry provides a concise exposition about a phenomenon and repeats key words multiple times. The findings suggested that exposure had an expected positive effect on both the speed of reading novel words and the scores in the vocabulary post-tests, both for L1 and L2 readers. In line with Ginestet et al. (2020) we find that the first exposure to the word is effortful while most subsequent exposures elicit a similar response time: we generalize this finding over L1 and L2 readers. Furthermore, even the least amount of exposure (2 novel word occurrences) was sufficient for the above-chance performance in semantic post-tests. The sufficient (above-chance) performance in the orthographic choice test required at least 4 exposures to the novel word in both L1 and L2 participants. These numbers are markedly smaller than the amount of exposure reported as sufficient in most previous studies (see the Introduction). We believe it is due to a tighter experimental control over our stimuli than in natural reading, which helped us reduce noise and increased the influence of word repetition on word learning.

Where the two cohorts of participants diverged was the number of exposures that led to maximum learning outcomes, within our experimental boundaries. L1 readers showed their maximum scores in orthographic choice, definition matching and definition prompting tasks after seeing the novel word 4 times; additional exposures did not improve their performance. Conversely, L2 readers improved their performance after 8 exposures to the novel word, as compared to 4.

It is tempting to describe this finding as an L1 advantage, stemming, say, from their larger vocabulary size and higher proficiency in the language of the test. Yet L1 readers reaching the optimal performance after 4 exposures was not a sign of the ceiling effect. This performance was far behind the maximum of any test scale. If anything, L2 participants were either on par with the L1 group in their scores or exceeded them. This is despite the fact that – besides the novel words – L1 readers are expected to have a better understanding of context, including superior knowledge of individual words in the passages, stronger probabilistic intuitions about collocations that these words form, and better understanding of the discourse structure, compared to L2 readers. Even though the limited knowledge of L2 is a major obstacle towards novel word learning (Elgort et al., 2018 and references therein), L2 readers out-perform native speakers in our tasks.

This consistent L2 advantage in novel word learning may be due to by the more extensive practice that L2 readers have in encountering and learning novel words in their foreign language, as opposed to native readers of their own language. The "multilingual advantage" has also been reported across many studies (see Cenoz, 2013 for a review) where experience with learning languages facilitated better performance than monolinguals on a variety of tasks (e.g., reading comprehension, vocabulary). In sum, the process of learning multiple languages may provide L2 readers with the tools to use cognitive resources more efficiently during vocabulary learning.

Another possible explanation for the observed L2 advantage stems from the format of our training materials, which implemented a so-called *massed* repetition of stimuli (each novel word occurrence was embedded into successive sentences and was present on the screen at the same time). This format is generally considered to be less conducive to learning than *spaced repetition*, with novel words spread apart from each other in different texts or time (see Cepeda et al., 2016 for a review).

Attention

Our novel manipulation of selective attention targeted its function that contributes to the preparation for action (Bradley 2003): unlike other functions, it has been relatively understudied in language learning research (see the Introduction). The present manipulation produced consistent expected results in the processing speed but not quality of learning (Figure 2 left). Awareness of the number of occurrences of a novel word through the Instruction manipulation led to significantly lower total fixations times to the first occurrence of that word in the conditions with 4 and 8 exposures, and to both occurrences of the word in the 2-exposures condition (Figure 2 left). This advantage disappeared in all subsequent occurrences of the novel word; thus, it was early and relatively short-lived.

Instruction did not appear to affect the accuracy of learning outcomes, as almost none of the vocabulary post-tests revealed a consistent influence of our manipulation. The impact of the Instruction condition was only found on semantic recognition performance in the L2 group: the readers in the no-instruction condition outperformed those in the instruction condition. We assumed that the presence of instructions would enable a reader to mobilize attentional resources for each learning opportunity more efficiently and consequently lead to better word learning. However, a reverse pattern was observed in semantic recognition performance of the L2 group. It is also noteworthy to recall here that, as cited in the Introduction, visual enhancement techniques aiming to increase learners' perception of targeted forms do not always guarantee better word learning (see e.g., Baleghizadeh et al., 2018; Izumi, 2002; Kang, 2003; Rott, 2007; Yang, 2004; Yeo, 2002). Since the duration of processing of passages or novel words did not differ between groups, the attention-inducing manipulation might lead readers to spend less amount of cognitive effort on context clues which are necessary to comprehend the text and to infer the meaning of novel words (Anderson, 1982; Laufer & Hulstijn, 2001). Also, considering their limited language proficiency, the L2 group is more likely to invest more mental effort for the process of the reading task. The presence of extra information about how many times a novel word would occur in the text, might increase the demand on the working memory capacity and limit their ability to gain maximum benefit from the text (see Schnotz & Kürschner, 2007 for a review). Still, the impact of instruction was observed in only one aspect of word knowledge and further studies are necessary to draw a more reliable conclusion. There was also no interaction between instruction and exposure in any vocabulary post-tests. However, the attention-inducing manipulation did influence learning speed, since both L1 and L2 groups recognized novel word spellings faster in the orthographic choice task when they received the instruction before reading. The impact of Instruction on RTs in the orthographic choice task was found in the first session only for the L1 group in both testing sessions for the L2 group. It should also be noted that neither the L1 nor the L2 group spent more time on reading the target novel words in the instruction condition. A similar finding was reported by Ginestet et al. (2020). Participants categorized as having high visual attention span recognized orthographic forms of novel words faster than those with lower visual span. Together this may suggest that delivering an attention-inducing instruction to a reader increased the amount of the attention resources allocated for word encoding and thus led to faster recognition of those words in the learning phase and faster recognition of their orthographic forms post-learning. This advantage warrants a further investigation of how to increase efficiency and quality of learning with a simple preamble to a very common object in everyday life of a reader, i.e., a text containing novel words.

The present study, along with several earlier reports (see the Introduction), demonstrates the benefits of studies that bring together L1 and L2 cohorts, online and offline tasks, and manipulations of multiple causal factors of word learning. While the present study has a tightly controlled factorial design, we believe that most progress in this research field will be made through combining studies like ours with more ecologically valid examination of natural reading.

References

Acheson, D.J., Wells, J.B., & MacDonald, M.C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40*(1), 278–289. https://doi.org/10.3758/BRM.40.1.278

Anderson, R. C. (1982). Allocation of attention during reading. In *Advances in psychology*, 8, 292–305.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77–117). Newark, DE: International Reading Association.

Baleghizadeh, S., Yazdanjoo, S., & Fallahpour, H. (2018). The effect of input enhancement on academic vocabulary learning among intermediate EFL learners in Iran. *TESL Reporter*, *50*(2).

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

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1–11.

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience, 117*(2), 369–380. https://doi.org/10.1037/0735-7044.117.2.369

Bransford, J. D., Franks, J. J., Morris, C. D., & Stein, B. S. (1979). Some general constraints on learning and memory research. *Levels of processing in human memory*, 331–354.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In *Eye movements in reading* (pp. 275–307). Academic Press.

Cenoz, J. (2013). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching*, *46*(1), 71–86. https://doi.org/10.1017/s0261444811000218

Cepeda, N., Pashler, H., Vul, E., Wixted, J., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Elgort, I., Brysbaert, M., Stevens, M., Assche, E. V. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, *40*, 341–366. https://doi.org/10.1017/S0272263117000109

Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2011). Predicting Robust Vocabulary Growth from Measures of Incremental Learning. *Scientific Studies of Reading*, 15(1), 71–91. https://doi.org/10.1080/10888438.2011.539076

Ginestet, E., Valdois, S., Diard, J., & Bosse, M. L. (2020). Orthographic learning of novel words in adults: effects of exposure and visual attention on eye movements. *Journal of Cognitive Psychology*, 1-20.

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., Yoon, H. (2017). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition, 21*(3), 1–22. https://doi:10.1017/S1366728917000219

Hulme, R. H., Barsky, D. & Rodd, J. M. (2019). Incidental learning and long-term retention of new word meanings from stories: The effect of number of exposures. *Language Learning: A Journal of Research in Language Studies. 69*(1), 18–43. https://doi.org/10.1111/lang.12313

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty, & M. H. Long (Eds.), The handbook of second language acquisition (pp. 349–381). (Blackwell handbooks in linguistics; No. 14). Malden, MA: Blackwell Publishing. https://doi.org/10.1002/9780470756492.ch12

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review, 61*(3), 355–382. https://doi.org/10.3138/cmlr.61.3.355

Hyde, T., & Jenkins, J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal Of Verbal Learning And Verbal Behavior*, *12*(5), 471–480. https://doi.org/10.1016/s0022-5371(73)80027-1

Izumi, S. (2002) Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*(4), 541–77. https://doi.org/10.1017/S0272263102004023

Joseph, H. S., Wonnacott, E., Forbes, P., & Nation, K. (2014). Becoming a written word: eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading. *Cognition*, 133(1), 238–248. https://doi.org/10.1016/j.cognition.2014.06.015

Kang, Y. (2003). Free voluntary reading: Cross-linguistic impact of parental reading behavior on children's English. *English Teaching, 58*, 49–60.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633. https://doi:10.3758/BRM.42.3.627

Kuznetsova, A., Brockhoff, B. & Christensen, H.B. (2013) lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of the lme4 package).

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied linguistics*, 22(1), 1-26. https://doi.org/10.1093/applin/22.1.1

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. R package version, Version: 1.5.2 Available at: https://CRAN.R-project.org/package=emmeans.

Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. Frontiers In Psychology, 6, 1171. https://doi.org/10.3389/fpsyg.2015.01171

Lowell, R., & Morris, R. (2017). Impact of contextual constraint on vocabulary acquisition in reading. Journal Of Cognitive Psychology, 29(5), 551–569. https://doi.org/10.1080/20445911.2017.1299155

Marian, V., Blumenfeld, H.K., & Kaushanskaya, M. (2007). Language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research, 50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Mohamed, A. A. (2018). Exposure frequency in l2 reading: An eye-movement perspective on incidental vocabulary learning. *Studies in Second Language Acquisition*, 1–25. https://doi:10.1017/S0272263117000092

Nation, P. (2001). Learning vocabulary in another language. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139524759

Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, *27*(1), 136–145.

Pashler, H., Johnston, J., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology,* 52, 629-651.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a Foreign Language*, *22*(1), 31–55.

Pellicer-Sánchez, A. (2016). Incidental l2 acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition, 38*(1), 97–130. https://doi.org/10.1017/S0272263115000224

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25–42.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Requin, J., Brener, J., & Ring, C. (1991). Preparation for action. In J. R. Jennings & M. G. H. Coles (Eds.), Wiley psychophysiology handbooks. Handbook of cognitive psychophysiology: Central and autonomic nervous system approaches (pp. 357–448). John Wiley & Sons.

Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning*, *57*(2), 165–199. https://doi.org/10.1111/j.1467-9922.2007.00406.x

Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, *6*, 72–78. https://doi.org/10.1016/0346-251X(78)90027-1

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual.* New York: Palgrave Macmillan. http://dx.doi.org/10.1057/9780230293977

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In *W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, Proceedings of CLaSIC 2010* (pp. 721–737). Singapore: National University of Singapore, Centre for Language Studies.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (Cambridge Applied Linguistics, pp. 3–32). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524780.003

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158. https://doi.org/10.1093/applin/11.2.129.

Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational psychology review*, 19(4), 469-508. https://doi.org/10.1007/s10648-007-9053-4

Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review*, 77, 406–418.

Song, Mi-Jeong. (2007). Getting learners' attention: Typographical input enhancement, output, and their combined effects. *English Teaching*, *62*(2), 193–215. https://doi.org/10.15858/engtea.62.2.200706.193

Teng, F. (2016). The effects of context and word exposure frequency on incidental vocabulary acquisition and retention through reading. *The Language Learning Journal*, *24*(2), 145–158.

Watanabe, Y. (1997). INPUT, INTAKE, AND RETENTION: Effects of Increased Processing on Incidental Learning of Foreign Language Vocabulary. *Studies in Second Language Acquisition*, *19*(3), 287–307. https://doi.org/10.1017/s027226319700301x

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, *20*(1), 232–245.

Winke, P.M., Godfroid, A., & Gass, S. M. (2013). Introduction to the special issue: Eye- movement recordings in second language research. *Studies in Second Language Acquisition*, *35.* 205–212.

Yang, C. (2004). Universal Grammar, statistics or both?. *Trends In Cognitive Sciences*, *8*(10), 451–456. https://doi.org/10.1016/j.tics.2004.08.006

Yeo, K. (2002). The effects of dictogloss: A technique of focus on form. *English Teaching*, *57*(1), 149–167.

Supplementary Materials

S1. Demographic information and tests of individual differences.

Individual differences measures: Participants were asked to fill out a language background questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007), a Reading Habits survey, and the Author Recognition Test (ART; Acheson, Wells, & MacDonald, 2008). LEAP-Q asks information about language proficiency, dominance, and preference, age of acquisition, exposure to each language, vision, hearing, language, and learning problems, and how many years of schooling they have. The Reading Habits survey asks the participant to self-rate (from 1-7) their reading skills by comparing their exposure to reading, understanding, and speed to other students their age. The ART instructs participants to check off recognized names from a list of real and foil authors, resulting in a score ranging from -65 to 65. The scores were calculated by subtracting the wrong guesses from the correct guesses. As confirmed by Welch's *t*-test, $t(41.75) = 2.07$, $p < .001$, the ART scores of the L1 ($M = 11.2$, SD = 5.9) outperformed those of the L2 ($M = 6.4$, SD= 9.9). The 95% confidence interval for the effect of language group on ART scores is between 0.11 and 9.3 percent.

At the beginning of the study, all participants completed a questionnaire rating their speaking, understanding, and reading skills in English from 1-10. L1 participants had mean scores above 9.6 for all these skills, and that L2 participants had a significantly lower mean and more variable scores (6.7 or above; all *p*s < 0.01) for all skills (see Table S1).

Table S1. Participant self-ratings in the English proficiency skills.

Skills	L1 Mean	L2 Mean
Speaking	9.87 (0.63)	6.75 (2.32)
Understanding	9.91 (0.42)	7.40 (2.22)
Reading	9.57 (0.81)	7.25 (1.89)

Note. Standard Deviations (SD) are presented in parentheses.

S2. Outputs of regression models

Table S2.Total novel word reading times (ms) in the L1 and L2

Instruction	Exposure	L1	L2
Yes	2	778.57 (408.33)	1154.52 (696.38)
No	2	894.23 (451.96)	1266.55 (620.15)
Yes	4	1374.48 (702.86)	1591.26 (767.64)
No	4	1574.19 (588.43)	1651.89 (603.64)
Yes	8	2355.55 (803.30)	3231.31 (1387.08)
No	8	2475.39 (735.05)	2868.67 (1305.18)

Note. Standard Deviations (SD) are presented in parentheses.

Table S3.ANOVA summary of a model fitted to total fixation time in L1 participants

	Sum Sq	Mean Sq	NumDF	DenDF	F value	p-value
Instruction	162689.40	162689.40	1.00	25.46	4.79	0.0379
Exposure	766251.59	383125.79	2.00	781.60	11.29	0.0000
Instruction: Exposure	221862.95	110931.48	2.00	782.10	3.27	0.0386

Note. Standard deviation of the by-subjects intercept is 43.84, standard deviation of word intercept is 14.84, and standard deviation of residuals is 184.22. N = 845 before trimming, N = 813 after trimming. $R^2 = 0.12$.

Table S4. ANOVA summary of a model fitted to total fixation time in L1 participants

	Sum Sq	Mean Sq	NumDF	DenDF	F value	p-value
Instruction	5824.38	5824.38	1.00	27.54	0.07	0.7886
Exposure	6086243.01	3043121.51	2.00	753.85	38.31	0.0000
Instruction: Exposure	226636.52	113318.26	2.00	754.21	1.43	0.2407

Note. Standard deviation of the by-subjects intercept is 110.08, standard deviation of word intercept is 8.98, and standard deviation of residuals is 281.82. N = 800 before trimming, N = 785 after trimming. $R^2 = 0.22$.

Table S5. Summary of the generalized linear mixed effect regression model fitted to the accuracy of orthographic choice task in L1 participants

	Estimate	SE	z-value	p-value
Intercept	-0.154	0.389	-0.394	0.693
Instruction=No	-0.036	0.363	-0.099	0.921
Exposure=4	0.959	0.279	3.435	0.001
Exposure=8	0.918	0.277	3.314	0.001
Session=2	0.847	0.229	3.696	0.000

Note. Standard deviation of the subject intercepts is 0.680, and the standard deviation of the word intercepts is 0.670. N = 450. R^2 = 0.24.

Table S6. Summary of the generalized linear mixed effect regression model fitted to the accuracy of orthographic choice task in L2 participants

	Estimate	Std. Error	z-value	p-value
(Intercept)	0.117	0.326	0.359	0.719
Instruction=No	-0.093	0.308	-0.304	0.761
Exposure=4	0.820	0.235	3.494	0.000
Exposure=8	1.397	0.251	5.556	0.000
Session=2	0.165	0.202	0.819	0.413

Note. Standard deviation of the subject intercepts is 0.605, and the standard deviation of the word intercepts is 0.565. N = 558. R^2 = 0.20.

Table S7. Summary of the generalized linear mixed effect regression model fitted to RTs of orthographic choice task in L1 participants

	Estimate	Std. Error	t-value	p-value
(Intercept)	1152.53	29.60	38.930	<0.000
Instruction=No	142.52	49.93	2.854	0.00431
Session=2	333.30	29.51	11.296	<0.000
Exposure=4	-60.18	40.70	-1.479	0.139
Exposure=8	-83.91	41.22	-2.036	0.0418
instructionNo:Session2	-267.95	33.08	-8.100	<0.000

Note. Standard deviation of the subject intercepts is 232.415, and the standard deviation of the word intercepts is 81.306. N = 450. R^2 = 0.18.

Table S8. Estimated marginal means contrasts of Instruction per level of Session for the L1 readers

Session	Contrast (Instruction)	Estimate	Std. Error	z-value	p-value
Session 1	Yes-No	143	49.9	2.854	0.004
Session 2	Yes-No	125	59.4	2.110	0.034

Table S9. Summary of the generalized linear mixed effect regression model fitted to RTs of orthographic choice task in L2 participants

	Estimate	Std. Error	t-value	p-value
(Intercept)	1309.43	35.01	37.403	<0.000
Instruction=No	189.15	38.02	4.975	<0.000
Exposure=4	-86.82	31.85	-2.726	0.006
Exposure=8	37.77	38.80	0.819	0.413
Session=2	70.81	30.16	2.348	0.019

Note. Standard deviation of the subject intercepts is 340.035, and the standard deviation of the word intercepts is 138.359. N = 558. R^2 = 0.18.

Table S10. Summary of the generalized linear mixed effect regression model fitted to definition matching scores in L1 participants

	Estimate	SE	z-value	p-value
Intercept	-2.178	0.494	-4.410	0.000
Session=2	-0.234	0.269	-0.872	0.383
Instruction=No	0.016	0.466	0.034	0.973
Exposure=4	0.827	0.351	2.355	0.019
Exposure=8	0.589	0.350	1.684	0.092

Note. Standard deviation of the random participant intercepts is 0.886, and the standard deviation of the random item intercepts is 0.691. N = 413. R^2 = 0.24

Table S11. Summary of the generalized linear mixed effect regression model fitted to definition matching scores in L2 participants

	Estimate	SE	z-value	p-value
Intercept	-2.376	0.514	-4.620	0.000
Session=S2	-0.443	0.264	-1.677	0.094
Instruction=No	1.088	0.503	2.163	0.031
Exposure=4	0.041	0.331	0.123	0.902
Exposure=8	0.544	0.318	1.710	0.087

Note. Standard deviation of the random participant intercepts is 1.04, and the standard deviation of the random item intercepts is 0.77. N = 459. $R^2 = 0.30$.

Table S12. Summary of the generalized linear mixed effect regression model fitted to definition prompting scores in L1 participants

	Estimate	SE	z-value	p-value
Intercept	-1.845	0.347	-5.319	0.000
Session=S2	0.159	0.165	0.963	0.336
Instruction=No	0.155	0.301	0.515	0.607
Exposure=4	0.549	0.211	2.595	0.009
Exposure=8	0.230	0.217	1.059	0.290

Note. Standard deviation of the random participant intercepts is 0.576, and the standard deviation of the random item intercepts is 0.596. N = 414. $R^2 = 0.24$.

Table S13. Summary of the generalized linear mixed effect regression model fitted to definition prompting scores in L2 participants

	Estimate	SE	z-value	p-value
(Intercept)	-1.760	0.325	-5.417	0.000
Session=2	0.264	0.162	1.628	0.103
Instruction=No	0.078	0.332	0.235	0.814
Exposure=4	0.013	0.212	0.061	0.951
Exposure=8	0.291	0.199	1.460	0.144

Note. Standard deviation of the random participant intercepts is 0.71, and the standard deviation of the random item intercepts is 0.46. N = 459. $R^2 = 0.23$.

S3: The paragraphs prepared for each level of Exposure condition

Aunith (Cell phone)

8 repetitions:
I dropped my *aunith* [3]on the ground, and now it is completely cracked. I quickly took my *aunith* to the store to get it repaired at a nearby mall. To fix my *aunith* will cost me over $100 which is a lot for me. From now on I promised I would keep my *aunith* in a case so that it doesn't break. Having my *aunith* with me every day is very convenient. I have a lot of reasons to use my *aunith* but games are my favourite. My grandpa bought a brand new *aunith* two weeks ago. He is usually very good at using it but still, sometimes, he finds an *aunith* hard to use and asks me for help.
The future of technology is more interesting than what we have right now.

4 repetitions:
I dropped my *aunith* on the ground, and now it is completely cracked. I quickly took my *aunith* to the store to get it repaired at a nearby mall. To fix my *aunith* will cost me over $100 which is a lot for me. From now on I promised I would keep my *aunith* in a case so that it doesn't break.
The future of technology is more interesting than what we have right now.
Kangaroos have large, long, and powerful legs. In a way, kangaroos can jump very high, sometimes three times their own height. Kangaroos like to fight for food, fun, land, and just about anything they can think of. Kangaroos are social animals and stay in groups of at least 3 or 4.

2 repetitions:
I dropped my *aunith* on the ground, and now it is completely cracked. I quickly took my *aunith* to the store to get it repaired at a nearby mall.
The future of technology is more interesting than what we have right now.
Kangaroos have large, long, and powerful legs. In a way, kangaroos can jump very high, sometimes three times their own height. Kangaroos like to fight for food, fun, land, and just about anything they can think of. Kangaroos are social animals and stay in groups of at least 3 or 4. Kangaroos are the only large animals which move by jumping. It is a very efficient way of motion because it preserves energy.

Ceammy (Lavender)

8 repetitions:
One can find and grow *ceammy* in almost all places. It is not surprising then that *ceammy* is very popular among gardeners who plant it every year. People like to put dried *ceammy* inside clothes to give them a nice air. Some people also like to use *ceammy* to perfume bed linen and closets. The unique, sweet smell of *ceammy* prevents mice, and other bad animals from the area. Typically, *ceammy* is also used in cake decorating, because it has a very nice look. The *ceammy* is

[3] The target words were italicized in this document for demonstration purposes only.

sometimes put in medicine, too. Mostly, *ceammy* has a purple color but there are also some blue ones.
The plants bring joy, health, and colour to wherever they are placed.

4 repetitions:
One can find and grow *ceammy* in almost all places. It is not surprising then that *ceammy* is very popular among gardeners who plant it every year. People like to put dried *ceammy* inside clothes to give them a nice air. Some people also like to use *ceammy* to perfume bed linen and closets.
The plants bring joy, health, and colour to wherever they are placed.
Planning an important event like a wedding can be very hard. There are many things to consider, like the location, food, and guests. Because there are lots of details to think about, people usually use professional organizations for it. This, of course, makes a wedding more expensive.

2 repetitions:
One can find and grow *ceammy* in almost all places. It is not surprising then that *ceammy* is very popular among gardeners who plant it every year.
The plants bring joy, health, and colour to wherever they are placed.
Planning an important event like a wedding can be very hard. There are many things to consider, like the location, food, and guests. Because there are lots of details to think about, people usually use professional organizations for it. This, of course, makes a wedding more expensive. However, most people agree that they are worth it. Weddings are a celebration of love and people remember their wedding forever.

Cruce (Hat)
8 repetitions:
A typical *cruce* is a head covering that people wear for many reasons. Some like to use a *cruce* for protection against rainy or cold weather. A stylish *cruce* also adds an elegant look as a fashion item. On the other hand, some people wear a *cruce* for an important event like graduation. In the past, a *cruce* was a display of social and economical statuses. In the military, a *cruce* may show rank. In many countries, police officers wear a *cruce* with a symbol of the department. The strong surface of a special *cruce* helps wearers stay safe while doing high risk jobs.
Some formal events only accept outfits with a single black colour.

4 repetitions:
A typical *cruce* is a head covering that people wear for many reasons. Some like to use a *cruce* for protection against rainy or cold weather. A stylish *cruce* also adds an elegant look as a fashion item. On the other hand, some people wear a *cruce* for an important event like graduation.
Some formal events only accept outfits with a single black colour.
When naming your new baby, it is important to think about the future. A unique name may seem like a good idea. However, a unique name may lead to bullying

in the future and because of this, your child may behave badly. They might also have a hard time getting a job at a later stage in life.

2 repetitions:
A typical *cruce* is a head covering that people wear for many reasons. Some like to use a *cruce* for protection against rainy or cold weather.
Some formal events only accept outfits with a single black colour.
When naming your new baby, it is important to think about the future. A unique name may seem like a good idea. However, a unique name may lead to bullying in the future and because of this, your child may behave badly. They might also have a hard time getting a job at a later stage in life. Celebrities often give their children crazy names like Apple and North. These names might not be as cute when these kids become adults.

Flyph (Mite)
8 repetitions:
You've probably never seen a *flyph* with the naked eye anywhere. However, just because you can't see them, that doesn't mean a *flyph* isn't around. Over 48,000 *flyph* species exist on Earth. Many live in the soil or water, but a *flyph* can also live on plants and animals like bees and mice. If you are seeing yellow marks on a leaf, a plant eating *flyph* might be the cause. There is also the hunting *flyph*, that actually eats the plant-eating kinds. Some *flyph* live on humans. However, the most disgusting type of *flyph* lives in houses, likes beds and causes allergies.
In cold places, some animals sleep during the coldest months of winter.

4 repetitions:
You've probably never seen a *flyph* with the naked eye anywhere. However, just because you can't see them, that doesn't mean a *flyph* isn't around. Over 48,000 *flyph* species exist on Earth. Many live in the soil or water, but a *flyph* can also live on plants and animals like bees and mice.
In cold places, some animals sleep during the coldest months of winter.
The recent changes in the economy makes an understanding of finance essential. Finance is a quickly growing field. It is studied by many students. These students must be knowledgeable in many subjects such as critical thinking, mathematical reasoning, and problem solving.

2 repetitions:
You've probably never seen a *flyph* with the naked eye anywhere. However, just because you can't see them, that doesn't mean a *flyph* isn't around.
In cold places, some animals sleep during the coldest months of winter.
The recent changes in the economy makes an understanding of finance essential. Finance is a quickly growing field. It is studied by many students. These students must be knowledgeable in many subjects such as critical thinking, mathematical reasoning, and problem solving. When these students graduate, there are several places they could work. For example, a bank or other companies on Wall street.

Mernt (Beagle, a kind of dog)
8 repetitions:
The exact origin of a *mernt* is unknown but these days they are very popular in the US. A typical *mernt* has a small to medium-sized body. Short legs and long, large, and soft ears give a *mernt* a cute look. The natural talents of the *mernt* are hunting, following, and finding. They can do this because a *mernt* has a very good sense of smell and a good memory. The widely known *mernt* has a coat of black, brown, and white colours. If you'd like to add a *mernt* to your family, remember that they are very loud creatures. So, a *mernt* may not be ideal for apartments. Many families with little children love having a pet in their home.

4 repetitions:
The exact origin of a *mernt* is unknown but these days they are very popular in the US. A typical *mernt* has a small to medium-sized body. Short legs and long, large, and soft ears give a *mernt* a cute look. The natural talents of the *mernt* are hunting, following, and finding.
Many families with little children love having a pet in their home.
Exercise is essential to a healthy lifestyle. Often, people think that a balanced diet is all that matters, but most of time this is not enough. Exercise allows people to burn extra calories, to build muscles, and to feel better about their body. There are different styles of exercise.

2 repetitions:
The exact origin of a *mernt* is unknown but these days they are very popular in the US. A typical *mernt* has a small to medium-sized body.
Many families with little children love having a pet in their home.
Exercise is essential to a healthy lifestyle. Often, people think that a balanced diet is all that matters, but most of time this is not enough. Exercise allows people to burn extra calories, to build muscles, and to feel better about their body. There are different styles of exercise. It could be indoors or outdoors, you can go running, swimming, road biking, golfing, hiking, and more. There is something for everyone.

Neak (Oven)
8 repetitions:
The original *neak* has been around for many thousands of years. The first traditional *neak* was a container placed on a fire to make a warm meal. The food was protected with leaves and put on top of the *neak* carefully. The *neak* was useful in early times for making simple food like bread. The first *neak* was invented by the Greeks. Today, more developed models of a *neak* can make many more things. Importantly, it holds heat for a long time, so the *neak* is very economical. These days, because it is necessary for our needs, a *neak* is very common to have in every home.
The invention of machines has changed the routines of life completely.

4 repetitions:
The original *neak* has been around for many thousands of years. The first traditional *neak* was a container placed on a fire to make a warm meal. The food was protected with leaves and put on top of the *neak* carefully. The *neak* was useful in early times for making simple food like bread.
The invention of machines has changed the routines of life completely.
In the middle ages, there was a large population decrease. Importantly, bad weather played a large part in that. There were many years where there wasn't enough rain, such that it killed the plants. A large number of people, both children and adults, didn't have enough food to eat.

2 repetitions:
The original *neak* has been around for many thousands of years. The first traditional *neak* was a container placed on a fire to make a warm meal.
The invention of machines has changed the routines of life completely.
In the middle ages, there was a large population decrease. Importantly, bad weather played a large part in that. There were many years where there wasn't enough rain, such that it killed the plants. A large number of people, both children and adults, didn't have enough food to eat. In addition to that, a great illness spread across the world. It is believed that around 35 million people died in Europe alone.

Plurk (Shovel)
8 repetitions:
To spend an afternoon outside, Nicole used a *plurk* to dig small holes in the garden. The dark brown handle of the *plurk* was made of a light wood. The smooth top part of the *plurk* was easy for her to hold. The sharp metal blade of the *plurk* helped Nicole complete her work really fast. She relied on the *plurk* to plant this space with different vegetable seeds. She admired how convenient the *plurk* was for her work. After she was done, she immediately put the *plurk* in the shed. This was because she did not want her *plurk* to get rusty in the backyard like last time.
Nicole always takes care of her belongings to avoid unwanted damage.

4 repetitions:
To spend an afternoon outside, Nicole used a *plurk* to dig small holes in the garden. The dark brown handle of the *plurk* was made of a light wood. The smooth top part of the *plurk* was easy for her to hold. The sharp metal blade of the *plurk* helped Nicole complete her work really fast.
Nicole always takes care of her belongings to avoid unwanted damage.
Throughout history, some cultures have been known for their skills in sailing. They are famous for travelling long distances from their home to unexplored places. They were also expert boat builders and have very powerful ships. They spent their lives either at sea or in battles.

2 repetitions:
To spend an afternoon outside, Nicole used a *plurk* to dig small holes in the garden. The dark brown handle of the *plurk* was made of a light wood.
Nicole always takes care of her belongings to avoid unwanted damage.
Throughout history, some cultures have been known for their skills in sailing. They are famous for travelling long distances from their home to unexplored places. They were also expert boat builders and have very powerful ships. They spent their lives either at sea or in battles. They did not protect themselves with strong shields. In fact, they actually went into battle wearing nothing on their head.

Rotch (Cha-cha-cha, a kind of dance)
8 repetitions:
The well-known *rotch* has an energetic rhythm and three quick steps. Moreover, the *rotch* involves lots of arm styling and sharp turns. A slower pace is better for learning, but the *rotch* can be performed at many speeds. Musically, the *rotch* was perhaps not a great innovation. However, the *rotch* became hugely popular because it was easy to learn. To study the *rotch* people went to Cuba. After they came back, they started to teach it to other people and the *rotch* became common at formal events. Today, the *rotch* is one of five accepted forms in competitions. The art is a powerful way to connect people of all ages and races.

4 repetitions:
The well-known *rotch* has an energetic rhythm and three quick steps. Moreover, the *rotch* involves lots of arm styling and sharp turns. A slower pace is better for learning, but the *rotch* can be performed at many speeds. Musically, the *rotch* was perhaps not a great innovation.
The art is a powerful way to connect people of all ages and races.
Research has discovered that talking, like sleep, is good for mental health. People often need therapy after traumatic events, such as loss of a loved one. Some people seek professional therapy immediately afterward. However, other people may avoid talking about their bad experiences.

2 repetitions:
The well-known *rotch* has an energetic rhythm and three quick steps. Moreover, the *rotch* involves lots of arm styling and sharp turns.
The art is a powerful way to connect people of all ages and races.
Research has discovered that talking, like sleep, is good for mental health. People often need therapy after traumatic events, such as loss of a loved one. Some people seek professional therapy immediately afterward. However, other people may avoid talking about their bad experiences. Recent studies show that even talking to a friend is an effective treatment. People should talk about the problems they want to solve.

Wurge (Strawberry)
8 repetitions:
When it is ready, a *wurge* is juicy and very sweet, and kids love it as a snack. A *wurge* is also very healthy because it has a lot of vitamin C. One picks a *wurge* with part of the stem still attached to the ground. Interestingly, a *wurge* does not continue to grow after being picked. There are different ways to eat a *wurge* such as fresh or in jam. My favourite way to enjoy a *wurge* is to use them for delicious drinks. In summer, my grandma likes to bake *wurge* pies with her classic recipe. Newspapers say that the average American eats at least one *wurge* a day.
Different local foods are available to people in each season of a year.

4 repetitions:
When it is ready, a *wurge* is juicy and very sweet, and kids love it as a snack. A *wurge* is also very healthy because it has a lot of vitamin C. One picks a *wurge* with part of the stem still attached to the ground. Interestingly, a *wurge* does not continue to grow after being picked.
Different local foods are available to people in each season of a year.
Recent studies show that people get sick more often in the winter months. One possible reason is that during the winter, people spend more time indoors with the windows closed. So, they have a higher risk of sharing air with someone who is sick. In the summer it is the opposite.

2 repetitions:
When it is ready, a *wurge* is juicy and very sweet, and kids love it as a snack. A *wurge* is also very healthy because it has a lot of vitamin C.
Different local foods are available to people in each season of a year.
Recent studies show that people get sick more often in the winter months. One possible reason is that during the winter, people spend more time indoors with the windows closed. So, they have a higher risk of sharing air with someone who is sick. In the summer it is the opposite. Another reason is the low level of sunlight, which is important for Vitamin D activation. Therefore, people buy more vitamins in the winter.

Conclusion

This thesis investigated the effect of morphological and identity priming in text reading in adult readers (L1 and L2) to answer the following questions: how does the morphological priming affect word recognition in natural reading (Chapter 2) and incidental word learning through reading (Chapter 3, only in L1 readers); and how do the joint effects of the number of word repetitions and attention to the repetitions affect novel word learning, and how many exposures to a novel word lead to sufficient and maximum semantic and orthographic knowledge (Chapter 4)?

The study of how words are learned and recognized can also offer invaluable insights into the nature of lexical representations. In visual word recognition, the repetition paradigm has long been used to understand how complex words are represented and processed in the mental lexicon. Repeated exposure to a novel word has also been found to be a strong predictor of word learning. Numerous studies have been conducted to determine the precise parameters of word repetition and how contextual (e.g., the informativeness of a text), lexical (e.g., length), or individual (e.g., age of learners) differences modulate this effect. Despite the fact that both lines of research (i.e., word learning and visual word recognition) have witnessed a great deal of interest in studying the effect of repetition from many perspectives, this thesis contribute to the existing literature by addressing research questions that had not previously been studied or had been explored only partially. The outcomes of this thesis will provide empirical evidence for theories the word learning and morphological processing.

The remainder of this chapter will include the research questions, a summary of how each chapter addresses these questions and the key findings of each chapter, a more in-depth discussion of the findings and the contributions of these findings to the existing body of knowledge, future study directions, and a conclusion.

1. *How does morphological relatedness (morphological priming) influence word recognition in natural reading of long texts?*

To address this question, Chapter 2 employed a specific method of priming, namely long-term inflection priming, in which prime-target pairs are separated by numerous intervening words. This priming was employed in L1 (Dutch and English) and L2 (English) long text reading, while identity priming was used as a baseline condition. The examination made use of the GECO database of the eye-movement data which includes eye movement records of monolinguals and bilinguals reading an entire novel (Cop, Dirix, Drieghe, & Duyck, 2017). Chapter 2 also examined whether the nature of prime-target pairs affects the presence and magnitude of priming by examining all pairwise combinations of base and inflected forms (e.g., *tree – tree*, *trees – trees*, *tree – trees*, *trees – tree*). Finally, Chapter 2 investigated how morphological priming affected different aspects of eye-movement behaviour and what cognitive processes are likely to underpin these effects by analysing several eye-tracking measures (i.e., total fixation time, gaze duration, first fixation duration, skipping rate, and regression rate).

Chapter 2 found little evidence of the presence of long-term morphological priming in natural text reading: The effect was only found in L2 reading between the inflected prime and the base target (e.g., *trees-tree*). Chapter 2 also reported ubiquitous evidence in favor of the identical priming condition which suggests that it is easier to *process* a word if the same form has been read earlier.

2. How does morphological priming affect the learning of novel words?

To answer this question, Chapter 3 investigated the efficiency of two learning mechanisms – i.e., decomposition and generalization – in native speakers of English. These mechanisms are presumed to facilitate the learning of novel words that are morphologically related to known words, using morphological knowledge and awareness abilities. All pairwise combinations of low-frequency bases (e.g., caltrop) and their novel derived forms (e.g., *caltroper*) were embedded into short stories as primes and targets. A *base-derived* prime-target set exemplified generalization (e.g., *caltrop-caltrop-CALTROPER*), while the *derived-base* prime-target set exemplified decomposition (e.g., *caltroper-caltroper-CALTROP*). The *base-base* and *derived-derived* identical sets served as controls. Learning of target words, which followed two repetitions of the prime, was measured by post-tests of orthographic and semantic learning to evaluate the efficiency of each mechanism. The efficiency of one mechanism over the other was expected to manifest itself in a difference in learning gains.

Chapter 3 found evidence in the advantage of decomposition (e.g., *caltroper-caltroper-CALTROP*) in orthographic and semantic learning outcomes. The benefit of identity priming has been observed in the learning of the novel words suggesting that it is easier to *learn* a word if the same form has been read earlier.

3. How do the effects of well-known word learning predictors– i.e., exposure and selective attention – interact during novel word learning for both L1 and L2 readers? How many exposures to a novel word lead to sufficient and maximum semantic and orthographic knowledge?

Chapter 4 investigated this issue in an incidental novel word learning paradigm by collecting eye-tracking data from L1 and L2 speakers of English. The effect of selective attention was examined by delivering attention-inducing instructions while the control group received no instructions. The effect of exposure was examined by presenting a novel word either 2, 4, or 8 times in short passages. Both online eye movement monitoring and the offline orthographic and semantic tests were used as indicators of word learning. *Sufficient* word learning was considered to be the minimal level of exposure to a novel word that led to above chance performance on the post-tests. *Maximum* word learning is defined as the level of exposure to a novel word after which point learning does not further improve in the framework of the study. To address the first objective, we investigated how the effects of selective attention and exposure interact either during the real-time learning phase or in the post-tests of orthographic and semantic knowledge. To address the second objective, the post-tests of each component of word knowledge, (i.e. orthographic recognition, semantic recognition, and semantic recall) were examined to determine which level of exposure to a novel word leads to *sufficient* and to *maximum* word learning.

Chapter 4 showed that word repetition is an effective predictor of word learning. L1 and L2 readers need about the same number of repetitions to obtain a sufficient level of word learning. However, L1 and L2 readers diverged on the number of exposures that led to maximum learning outcomes. A higher number of repetitions of a novel word is more likely to result better word learning in L2, whereas L1 readers can reach their maximum learning with fewer repetitions of a novel word. On the other hand, Chapter 4 found that the attention-inducing manipulation of instruction led to faster recognition of novel words in both groups, but only led to a speed-up in eye-movements for the L1 readers.

The findings will be discussed in further detail below, as well as their addition to the current body of knowledge.

Long-term morphological priming might help word recognition, but only in L2

Chapter 2 reported an advantage of base forms primed by inflected forms (e.g., *trees-tree*) over unprimed base forms in both gaze duration (10 ms) and total fixation duration (15 ms) measures in L2 reading. Importantly, Chapter 2 found that this effect was limited to L2 reading and no other reliable statistical evidence was found in L1 reading (English or Dutch) for either type of morphological priming under consideration. This finding offers a new insight into morphological processing literature since it is contrary to the findings of many single-word priming studies (e.g., Münte, Say, Clahsen, Schiltz, & Kutas, 1999; Rodriguez-Fornells, Münte, & Clahsen, 2002; Raveh & Rueckl, 2000; Royle, Drury, Bourguignon, & Steinhauer, 2012; Rueckl & Aichler, 2008; Rueckl, Mikolinski, Raveh, Miner, & Mars, 1997; Stanners et al., 1979; Weyerts, Münte, Smid, & Heinze, 1996).

The null effects in L1 reading raises concerns about the generalizability of current theories of morphological processing. As discussed in the Introduction, theories of morphological processing and representation propose every possible scenario for the organization and functioning of the mental lexicon, yet they mostly agree on the facilitatory effect of morphological priming on the recognition of any given word (Butterworth 1983; Taft & Forster, 1975; Caramazza, Laudanna, & Romani, 1988; Schreuder & Baayen, 1995; Giraudo & Grainger, 2000; Grainger, Colé, & Segui, 1991; Plaut & Gonnerman, 2000; Rueckl & Raveh, 1999; Baayen et al., 2011). The morphological priming effect has been so repeatedly observed in L1 visual word recognition research to the point that Amenta and Crepaldi (2012) state in their review that doubting its existence is pointless given the amount of evidence.

Notably, the vast majority of empirical evidence underpinning the theoretical accounts of morphological processing, and so the inference of Amenta and Crepaldi (2012), comes from single word recognition paradigms. By introducing a methodological twist (by employing long-term morphological priming in natural reading of long texts) to the traditional approach, Chapter 2 nullifies the previously reported priming effect. Furthermore, Chapter 2 questions the ecological validity of isolated word recognition paradigms and highlights the importance of revising existing theories in order to account for behaviours observed outside the demands and artificial constraints of laboratory tasks.

Chapter 2 made use of the eye-tracking data collected while reading a complete novel in order to investigate morphological priming in long text reading. Eye-tracking offers a reliable way to examine naturally occurring reading behavior (Rayner, 1998). In the morphological processing literature, eye tracking is mostly used to study complex word processing in the context of sentence reading (with a few exceptions of studies that present words in isolation, e.g., Kuperman et al., 2009; Hyönä, Laine, & Niemi, 1995; see Bertram, 2011 for a review). The use of eye-tracking in long text reading however is extremely rare (e.g., Kamienkowski et al., 2018). Although sentence reading paradigms provide a more natural setting than isolated word recognition paradigms, their ecological validity remains limited as these sentences are carefully constructed to highlight the stimuli of interest and do not have cohesive ties to the other sentences in the stimulus list. This study, along with Kamienkowski et al. (2018), offers a new way to study the role of morphological priming in reading by means of eye tracking data collected during reading a long text.

Chapter 2 involved a comparison of L1 and L2 readers which provides a new perspective to the existing discussion. Chapter 2 reported that the long-term inflectional priming has only been found in L2 readers. This result contradicts previous findings in the morphological processing literature. To be more precise, a review of the literature on L2 morphological priming reveals two opposing viewpoints on this matter: Some argue that the morphological priming effect exists only in L1 reading but not in L2 (e.g., Clahsen & Neubauer, 2010; Jacob, Fleischhauer, & Clahsen, 2013; Jacob et al., 2017; Kirkici & Clahsen, 2013; Neubauer & Clahsen, 2009; Silva & Clahsen, 2008), whereas others argue that it exists in both L1 and L2 reading (Basnight-Brown et al., 2007; Coughlin & Tremblay, 2014; De Grauwe et al., 2014; Diependaele et al., 2011; Feldman et al., 2010; Foote, 2015). Although the finding of Chapter 2 contradicts the existing evidence, Chapter 2 argues that this discrepancy may be due to statistical reasons. For example, L2 readers read at a slower rate, compared to the L1, therefore any effect on reading times would be larger and easier to detect in statistical models for L2 readers (see Jacob, 2018). To the best of our knowledge, this is the first study that examines the morphological priming effect in L2 long text reading. Therefore, to fully comprehend the nature of this effect, Chapter 2 encourages additional studies on the topic.

Morphological priming might help word learning in L1, but only in the decomposition (derived-derived-BASE) condition

Generalization and decomposition are two potential learning mechanisms involved in inferring the meaning of an unknown word based on existing word and morphology knowledge: The ability to infer the meaning of a (unknown) derived word (e.g. *contactless*) by computing the meaning of the known base (e.g., *contact*) and the derivational morpheme (e.g., -*less*) is referred to as generalization. The ability to infer the (unfamiliar) meaning of the base (e.g., *contact*) given the knowledge of the complex word (e.g., *contactless*) and the derivational morpheme (-*less*) is referred to as decomposition.

As a result of comparing the efficiency of generalization (e.g. *caltrop-caltrop-CALTROPER*) and decomposition (e.g., *caltroper-caltroper-CALTROP*) mechanisms, Chapter 3 discovered that the decomposition mechanism is a stronger, more efficient mechanism than generalization for both orthographic and semantic learning. In the decomposition condition, Chapter 3 reported a speed advantage to orthographic choice response times. The efficiency of the decomposition mechanism is also observed in the analyses of errors in the definition prompting task. This analysis revealed that when prompted for definitions, participants defaulted to semantic decomposition as opposed to providing a definition for a complex word as required by the question. Taken together, these findings suggest that morphological knowledge and awareness might help word learning but in only one way: namely, through semantic decomposition. The findings from this study make several contributions to the current literature outlined below.

To our knowledge this is the first study that directly compares the efficiency of decomposition and generalization as learning mechanisms. In word learning literature, the role of generalization has been studied more often (e.g., Bertram, Laine, & Virkkala, 2000; Carlisle, 2000; Mahony et al., 2000; Wood, Mustian, & Cooke, 2012; Wysocki & Jenkins, 1987), whereas our understanding on the role of decomposition is still limited to a few studies (e.g., Dawson et al., 2021; Ginestet et al., 2021; Pacton et al., 2013; Tucker et al., 2016). Nevertheless, this study demonstrated that this under studied mechanism, can be a valuable tool for learning the form and meaning of unfamiliar words.

The second contribution of this study is that it provided a deeper insight into the efficiency of the decomposition mechanism by examining both the breadth (orthographic and semantic components) and depth (recognition and recall tasks) of word knowledge. The existing studies of decomposition primarily concentrate only on a single component of word knowledge (e.g., Pacton et al., 2013; Ginestet et al., 2021). This study examines the issue from a broader perspective, as befits the multidimensional nature of words.

Moreover, another significant contribution of this research is that it investigates the morphological priming effect in prime-target pairs in both directions: from base to derived and from derived to base. Even in the morphological processing literature, where priming manipulation is extremely common, there are only a few studies that directly compare the magnitude of priming elicited in prime-target pairs of the base-derived and derived-base conditions. In this respect, this study addresses a significant gap, not only in the word learning literature, but also in the morphological processing literature. Marslen-Wilson et al. (2008) is one of these few studies which compared the roles of decomposition and generalization in a masked priming task. They reported that decomposition provides a processing advantage over generalization, which is consistent with the findings in Chapter 3. From the word learning perspective, Chapter 3 argues that the advantage of semantic decomposition could be related to the amount of information provided by a base vs a derived form. A transparent derived form (e.g., *farmer* as *farm* and *-er*) physically includes the base and allows a learner to extract base-related information with each occurrence, whereas the base form does not provide such a direct benefit to its derived form. However, more research in both fields is needed to determine what other factors may play a role in this effect. To sum up, this study fills a gap in the literature, and encourages more direct comparisons of *base-derived* and *derived-base* prime-target pairs in the fields of morphological processing and word learning

The finding showing the efficiency of decomposition as a learning mechanism has important implications for applied work on morphological instruction and vocabulary acquisition. Learning the meanings of simple words through semantic decomposition can be a useful study technique. This suggests that, in instructional material, for example, a learner will benefit the most if they encounter the complex form first and the simplex form second. Additionally, reading materials that adolescents and adults tend to encounter contain a high number of low frequency words with familiar affixes and unfamiliar bases (Nagy & Anderson, 1984). Thus, making these unfamiliar base forms available after reading training would improve readers chances of learning.

Identical Repetition helps the processing and learning of a (novel) word

Identity priming was utilized as a baseline in both Chapters 2 and 3 and was found to have clear effects: Words primed by their identical forms are recognized faster (Chapter 2) and learned better (Chapter 3).

Chapter 2 reported a facilitatory effect of identity priming on eye-movement measures of word recognition in long text reading, despite most of the repeated words being separated by hundreds of words. This effect has been observed in all samples (L1 English and Dutch, and L2 English) and across multiple eye-movement measures (total fixation times and regression rates in English L1 and L2 reading, and regression and skipping rates in Dutch L1 reading). These results are similar to those of Kamienkowski et al. (2018), who found that identity priming has a facilitative effect on Spanish data. Whereas Kamienkowski et al. (2018) found identity priming effects appeared in early reading measures, Chapter 2 mostly found these effects in late measures (except for the effect in skipping as found for Dutch L1). One possible explanation of this difference could be attributed to the different degrees of orthographic transparency of the examined languages. When the degree of transparency is high, such as in Spanish, the effects can be found in early measures, and when the degree of transparency is low, such as in English, the effects can be found in late measures. Dutch is in the middle of this spectrum, and contains evidence from both measures. In visual word recognition literature, repetition priming has been examined for decades (e.g., Forbach et al., 1974). Prior exposure to a lexical item is now widely acknowledged to have a facilitatory effect on its subsequent occurrences. However, Chapter 2 is one of the few studies that extends the existing literature to the natural reading of long texts.

Chapter 3 reported the learning advantage of words in the identity priming condition over those in the morphological priming condition. When the same form is presented repeatedly, they are learned better, as evidenced by higher semantic and orthographic learning accuracy outcomes. They are also retrieved more easily, as evidenced by faster recognition of the target word's correct spelling. Chapter 4 further investigates the influence of word repetition on learning by varying the number of repetitions, allowing for a more systematic comparison of the effect.

Taken together, these findings suggest that repetition of the same form of the word would be more likely to improve quality of learning (base-base and derived-derived, Chapter 3) while lowering processing effort (base-base and inflected-inflected, Chapter 2). These findings contribute to the body of knowledge showing the effectiveness of repetition in both word recognition and word learning literature. These findings also offer a practical application for vocabulary teaching by emphasizing the importance of exposure to the identical form rather than the morphologically related form in the first few exposures for greater learning gains.

Still, not all repetitions are the same for L1 and L2, and for semantic and orthographic learning.

One of the aims of Chapter 4 was to discover what level of exposure to a novel word (2, 4, or 8 times) would result in sufficient word learning and what level of exposure would result in the maximum word learning within the boundaries of the study's framework.

In incidental word learning, repeated exposure to a word is considered to be a key predictor of word learning. However, the precise parameters of this effect in L1 and L2 learning is not fully understood. For example, studies have reported varying answers to the question of what level of exposure to a novel word is required for the *sufficient* and *maximum* vocabulary learning. For instance, Hulme et al. (2019) proposed that two exposures to a novel word would be sufficient for semantic learning, though semantic learning would still be improved after eight exposures in L1 reading. On the other hand, to Godfroid et al., (2017), sufficient semantic knowledge would require around 8-10 exposures to a novel word, and an L1 reader can continue to improve their learning even after 23 exposures to the word. In L2 studies, the variation is even more pronounced, ranging from six exposures (Rott, 1999) to 20 exposures (Waring & Takaki, 2003).

Chapter 4 discovered that while two exposures to a novel word were sufficient for above-chance performance in the semantic post-tests, it took at least four exposures to achieve this performance in the orthographic post-test. This pattern was observed both in L1 and L2 readers. Perhaps the most intriguing finding of Chapter 4 was that L2 readers continued to benefit from additional exposures and reached their maximum learning after the 8-exposures condition in orthographic choice, definition matching and definition prompting tests. On the other hand, L1 readers reached their maximum learning after four exposures to a novel word and showed no progress with additional exposures in the eight-exposure condition in either of the tests. It should be noted that L1 readers' test scores were never close to the maximum on any test scale, which indicates that ceiling effects were not the reason why L1 readers did not benefit from the extra repetition. In addition, we should note that L2 readers perform at least as well as L1 readers in each test. Taken together, these findings may imply that L2 readers have a greater capacity for learning novel words. Chapter 4 discusses how factors like knowing multiple languages and being open to learning new words could explain the advantage for L2 readers in novel word learning. Still, more research should be conducted with a systematic comparison of L1 and L2 learners and of different aspects of word knowledge to gain a better understanding of the issue.

The reading times for each individual occurrence of the novel words were also examined. One of the most common findings in eye tracking studies of word learning is that reading novel words for the first time takes longer, but with more exposures to the word, reading times decrease and eventually reach an optimum speed (see a review by Winke, Godfroid, & Gass, 2013). Consistent with these findings, Chapter 4 reported that novel word reading times for L1 readers decreased immediately following the first exposure, whereas reading times for L2 readers decreased gradually and plateaued after 3-4 exposures. This was also consistent with previous research that observed similar trends among L1 and L2 readers (e.g., Pellicer-Sánchez, 2016; Godfroid et al., 2017; Ginestet et al., 2020).

To sum up, this study expanded previous research by including both online and offline (i.e. orthographic and semantic components) measures of word learning in L1 and L2 readers of English to determine where these effects occur during and after reading. Chapter 4 indicates word repetition had the expected positive effect on both learning and eye-tracking measures and in both the L1 and L2 groups. However, more repetition did not necessarily enhance orthographic and semantic word knowledge as observed in L1 readers.

These findings also offer valuable insights to the language teaching domain. Presenting a word at least four times could be a beneficial instructional strategy, as both form and semantic learning would develop for skilled and less skilled speakers of the language. The results also suggest that form learning requires more effort than semantic learning for both L1 and L2 learners. Therefore, implementing form-focused practices into instructional contexts can help to speed up the learning of the orthographic knowledge especially in the early stages of learning. These findings also suggest that L2 learners can learn novel words just as well as L1 learners, although L2 learners will require more exposure to a novel word. This can be especially critical in instructional settings with a mix of L1 and L2 learners, where the content includes novel words, such as academic terms, that are unfamiliar to both groups.

Attention facilitates processing speed in L1 and the quality of orthographic learning in both L1 and L2 groups.

Chapter 4 additionally investigated whether selective attention plays a unique role in novel word learning and whether it modulates the impact of exposure. For this purpose, half of the readers both in the L1 and the L2 groups received attention-inducing instructions about the number of occurrences of the novel word in the upcoming text, while the other half did not. Chapter 4 found that attention-inducing instructions increased learning quality of orthographic learning for both groups and the processing speed in L1 readers.

The study reported that the attention-inducing manipulation could only account for the decrease in reading times up to first two occurrences of a novel word in L1 reading, while L2 reading times did not benefit from the instruction manipulation. These results suggests that the instruction manipulation led to an early and short-lived effect for L1 readers but did not influence L2 readers.

Chapter 4 also reported that the instruction manipulation improved the quality of learning for the orthographic component of word knowledge. That is, both groups of readers recognized the correct spelling of a novel word faster when they were informed about how many times the novel word would be presented in the text they were reading.

Finally, Chapter 4 reported that the instruction manipulation had an effect on the semantic recognition performance of the L2 group. L2 readers who did not receive the attention-inducing manipulation performed better than those who did.

To our knowledge, Chapter 4 is the first study to examine the effects of attention and repetition, together, in an incidental word learning task. In this regard, these findings contribute to the field of word learning research by demonstrating the distinctive role of exposure in word learning, compared to the limited impact of attention. One study that could be relevant to ours was conducted by Ginestet et al. (2020). Ginestet et al. did not utilize any kind of manipulation to guide the participants' attention; instead, they categorized readers as having a high or poor visual attention span. The researchers then investigated how the number of exposures to isolated novel forms (one, three, or five times), influences the orthographic learning of these forms. The Ginestet et al. (2020) findings resemble those of Chapter 4 in several ways.

First, Ginestet et al. (2020) observed that in the first few repetitions, the reading time of participants categorized as having high visual attention span showed a sharper decrease. This seems to be comparable to what we discovered in our research. This suggests that the L1 group in our study that received the attention instruction may have used their attentional resources more strategically, resulting in a considerable decease in reading times, especially for the first few exposures to the novel words. On the other hand, reading in a second language is a more challenging process that may require L2 readers to commit all of their available attentional resources to digesting text. As a result, they are unable to process a second piece of information provided by the instruction, or have severely limited resources to do so. This might also explain why L2 readers who received the instruction performed poorly in the semantic recognition test. It is possible that the instruction increased their processing load, thus impacting the development of the form-meaning link.

A second similarity is that Ginestet et al. (2020) reported a similar trend in the outcomes of the orthographic recognition task. In Ginestet et al. (2020) the correct spelling of a novel word was recognized faster by readers with a high visual attention span than by those with a low visual attention span. Notably, neither group's accuracy scores were higher than the other. In Chapter 4, we found that attention-inducing manipulation led to faster recognition of the novel words while having no effect on accuracy scores. Therefore, based on Ginestet et al., we propose that attention-inducing manipulation may allow both L1 and L2 readers to employ their attentional resources more efficiently to encode novel words, resulting in higher-quality representations in the lexicon.

The key novelty of Chapter 4 was that it used an incidental word learning paradigm to investigate two important predictors of word learning: attention and exposure. Schmidt (2001) claims that learning can only emerge with attention and noticing the word. However, the outcomes of this study do not fully support this view. Our findings revealed that while paying attention to novel forms is critical for form learning, it has almost no effect on semantic learning. On the other hand, we discovered that word repetition is a significant predictor of incidental word learning and influences every aspect of word knowledge.

This research is also significant from a practical standpoint. In instructional settings, for example, educators can deliver instructions that will prepare students for the task. Even if there is no improvement in learners' exam scores, it can still have a positive impact on other dimensions of learning quality, such as quicker recall or identification.

Chapter 4 offers a comprehensive and systematic investigation of the effects of exposure and attention on vocabulary learning by combining multiple learning outcomes (accuracy, RTs, reading durations), different word components (orthographic and semantic), and two populations (L1 and L2 readers).

Limitations and Future Work

Chapter 2

To recap, Chapter 2 made use of the GECO database of eye movements recorded while reading a novel for comprehension. Although the investigation of natural texts has higher ecological validity than the study of artificial stimuli, it nevertheless still has limitations. Tightly controlled studies have greater control over the stimuli, allowing them to specify the features of lexical items as well as the number of intervening items. Although these artificial arrangements usually violate ecological validity, they do also eliminate sources of variance. To reduce undesirable variance in Chapter 2, outlier stimuli were excluded, and statistical considerations were applied. Furthermore, the consistency of the results across Spanish, Dutch, and English corpora implies that the null effects of long-term inflection priming in L1 reading is not a coincidence. However, further investigation is necessary in order to achieve a more definitive conclusion.

In Chapter 2, inflection priming in regular nominal and verbal forms was chosen as it was most likely to elicit effects due to a considerable orthographic and semantic overlap between the base and inflected forms. Similar research should be also conducted on irregular inflection, as well as derivation and compounding considering their semantic features. Such a comparison may be intriguing for L2 readers in particular, given that we observed long-term morphological priming in this group.

Chapter 2 did not look at prime-target pairs that are only orthographically connected (for example, scandal – scan). However, studies suggest that L2 speakers are more likely than L1 speakers to rely on surface features of a word in reading. As a result, this sort of comparison might indicate whether long-term morphological priming in L2 is attributable to morphological overlap or purely to overlap in form.

Chapter 3

In Chapter 3, we found the base forms primed by the inflected forms were recognized faster. Future studies should also address the effect of orthographic priming on novel word learning (e.g., *scan* and *scandal*). A condition like this would enable to examine whether this advantage can be attributed to solely the orthographic overlap between a prime and target.

Chapter 3 should be expanded to investigate less proficient readers as well due to less experienced readers being more likely to be unfamiliar with base forms and derivational morphemes which constitute the derived form. Similarly, their ability to infer the meaning of a base form from a complex word through decomposition would be limited (see Anglin, 1993; also see Carlisle, 2010 for a review). Therefore, the observed decomposition advantage reported in Chapter 3 may not be seen in this group. In addition, Chapter 3 mainly considered pairs of words with highly transparent morphological relationships (e.g., *caltrop-caltroper* and *gimbal-gimbalist*). Since the semantic relationship between these lexical items is more noticeable, learning the base through decomposition may be more achievable. Future studies should also investigate word pairs with less transparent morphological relationships (e.g., *depart-department*) to determine to what extent transparency influences the efficiency of decomposition.

Chapter 4

Chapter 4 found that, the group that received the attention-inducing instruction showed similar reading and learning behaviours as the high visual attention span group in Ginestet et al. (2020). This study might be reproduced in the future by taking into account the participants' visual attention span, which could improve our understanding of the function of attention in novel word learning. Future studies could additionally investigate alternative methods of manipulating selective attention to novel words. Chapter 4 only found a limited effect of the attention-inducing instruction with a short-lived effect on reading times (in L1) and faster recognition of the orthographic forms (in L1 and L2). Critically, the attention-inducing manipulation did not lead to higher accuracy scores as suggested by Schmidt (2001). The question then becomes whether this limited effect of attention is due to the way we manipulate selective attention to novel words, or the effect of repetition (even if it is as few as two) is so robust that it does shadow any influence of attention. In Chapter 4, we chose to manipulate selective attention to occurrences of novel words by presenting instructions concerning the number of exposures to some participants. To answer this question, future studies should consider other ways to draw readers' selective attention to occurrences of novel words (as introduced in Chapter 4, e.g., Baleghizadeh, Yazdanjoo, & Fallahpour, 2018; Song, 2007; Izumi, 2002; Watanabe, 1997; also see Hulstijn, 2003 for a review).

In Chapter 4, we defined maximum word learning by taking into account the number of novel words that were repeated (2,4, or 8 times) during the study. As a result, it is possible that L1 learners' performance will increase with further exposure to the novel words outside of the study's parameters. Similarly, it is uncertain whether the L2 group's current level of learning would be maintained or improved if the novel words were repeated more than 8 times. Importantly, neither group's performance in the post-test demonstrated the presence of the ceiling effect. As a result, future research should look at the impact of word repetition using a broader range of repetition to see how the performance of both groups changes with more exposures.

Finally, Chapter 4 examined incidental word learning through reading of short passages designed for the purpose of the experiment. This type of text provides more control over the presentation of stimuli. Also, it helps to reduce variances which natural texts may have (e.g., the length of text, the amount of information provided). However, texts designed for experimental purposes may also have drawbacks such that they might include structures that are not seen frequently in natural reading (e.g., presenting two separate stories in a single passage). Therefore, future research should examine the joint effects of selective attention and exposure on word learning in less constrained reading contexts. For example, it could be a novel that contains foreign or technical words that are unfamiliar to readers (e.g., Godfroid et al., 2017; Saragi et al., 1978).

Conclusions

This thesis examines the influence of morphological and identity priming from a variety of angles to understand what aspects of repetition influences word recognition and novel word learning in adult first (L1) and second (L2) readers. The findings revealed that repeated exposure to identical novel words improves the quality and likelihood of learning them in incidental learning tasks, and increases their recognition speed in natural reading. Yet, as discussed in Chapter 4, the number of exposures to a novel word affect L1 and L2 readers' word learning performance differently. L1 readers hit a learning plateau after 4 exposures, therefore increased exposures does not necessarily result in continuously better word learning. More importantly, this thesis highlights that there is an asymmetry in morphological priming effects: the facilitatory effect appears only when a morphologically complex word is followed by its base form (e.g., Chapter 2:*trees-tree* and Chapter 3:*caltroper-caltrop*). This impact manifests itself as a word recognition advantage in L2 readers (Chapter 2) and a word learning advantage in L1 readers (Chapter 3).

To sum up, the goal of this thesis is to provide empirical evidence for mental lexicon research through a series of reading studies that include both experimental and corpus-based research. This thesis offers the advantage of examining what aspects of morphological and identity priming are influential by bringing together two lines of work (text reading and word learning), different populations (L1 and L2), different languages (Dutch and English), and online (reading by means of eye tracking) and offline tasks (post-tests of orthographic and semantic learning). The findings of this study can also aid language instructors in developing effective word-learning strategies.

References

Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology*, *3*, 232.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society of Research in Child Development, 58*(10, Serial No. 238).

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*(3), 438.

Baleghizadeh, S., Yazdanjoo, S., & Fallahpour, H. (2018). The effect of input enhancement on academic vocabulary learning among intermediate EFL learners in Iran. *TESL Reporter*, *50*(2).

Bertram, R., Laine, M., & Virkkala, M. M. (2000). The role of derivational morphology in vocabulary acquisition: Get by with a little help from my morpheme friends. *Scandinavian Journal of Psychology*, *41*(4), 287-296.

Bertram, R. (2011). Eye movements and morphological processing in reading. *The Mental Lexicon*, *6*(1), 83-109.

Basnight-Brown, D., Chen, L., Hua, S., Kostić, A. and Feldman, L. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of Memory and Language*, 57(1), pp.65-80.

Butterworth, B. 1989. Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process.* Cambridge, MA: MIT Press.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*(3), 297-332.

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading research quarterly*, *45*(4), 464-487.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, *12*(3), 169-190.

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

Coughlin, C. And Tremblay, A. (2014). Morphological decomposition in native and non-native French speakers. *Bilingualism: Language and Cognition*, 18(3), pp.524-542.

Cop, U., Dirix, N., Van Assche, E., Drieghe, D., & Duyck, W. (2017). Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism: Language and Cognition*, *20*(4), 747-769.

Dawson, N., Rastle, K., & Ricketts, J., 2021. Bridging form and meaning: support from derivational suffixes in word learning. *Journal of Research in Reading*, 44(1), pp.27-50.

De Grauwe, S., Willems, R. M., Rueschemeyer, S.-A., Lemhöfer, K., and Schriefers, H. (2014). Embodied language in first- and second-language speakers: neural correlates of processing motor verbs. *Neuropsychologia* 56, 334–349. doi: 10.1016/j.neuropsychologia.2014.02.00

Diependaele, K., Grainger, J., & Sandra, D. (2012). Derivational morphology and skilled reading. *Cambridge Handbook of Psycholinguistics*, 311-332.

Feldman, L. B., Kostić, A., Basnight-Brown, D. M., Durđević, D. F., & Pastizzo, M. J. (2010). Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism (Cambridge, England)*, *13*, 119–135. https://doi.org/10.1017/S1366728909990459

Foote, R. (2015). The Storage And Processing Of Morphologically Complex Words In L2 Spanish. *Studies in Second Language Acquisition*, 39(4), 735-767.

Forbach, G. B., Stanners, R. F., & Hochhaus, L. (1974). Repetition and practice effects in a lexical decision task. *Memory & Cognition*, *2*(2), 337-339

Ginestet, E., Shadbolt, J., Tucker, R., Bosse, M. L., & Deacon, S. H. (2021). Orthographic learning and transfer of complex words: insights from eye tracking during reading and learning tasks. *Journal of Research in Reading*, *44*(1), 51-69.

Ginestet, E., Valdois, S., Diard, J., & Bosse, M. L. (2020). Orthographic learning of novel words in adults: effects of exposure and visual attention on eye movements. *Journal of Cognitive Psychology*, 1-20.

Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and cognitive processes*, *15*(4-5), 421-444.

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, A., Yoon, H. (2017). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition*, *21*(3), 1–22. https://doi:10.1017/S1366728917000219

Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of memory and language, 30*(3), 370-384.

Hyönä, J., Laine, M., & Niemi, J. (1995). Effects of a word's morphological complexity on readers' eye fixation patterns. In *Studies in visual information processing* (Vol. 6, pp. 445-452). North-Holland.

Hulme, R. C., Barsky, D., & Rodd, J. M. (2019). Incidental learning and long-term retention of new word meanings from stories: The effect of number of exposures. *Language Learning*, *69*(1), 18-43.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty, & M. H. Long (Eds.), The handbook of second language acquisition (pp. 349–381). (Blackwell handbooks in linguistics; No. 14). Malden, MA: Blackwell Publishing. https://doi.org/10.1002/9780470756492.ch12

Izumi, S. (2002) Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, *24*(4), 541–77. https://doi.org/10.1017/S0272263102004023

Jacob, G., Heyer, V. and Veríssimo, J. (2017). Aiming at the same target: A masked priming study directly comparing derivation and inflection in the second language. *International Journal of Bilingualism*, 22(6), 619-637.

Jacob, G., Fleischhauer, E. And Clahsen, H. (2013). Allomorphy and affixation in morphological processing: A cross-modal priming study with late bilinguals. *Bilingualism: Language and Cognition*, 16(4), 924-933.

Jacob, G. (2018). Morphological priming in bilingualism research. *Bilingualism: Language and Cognition*, *21*(3), 443-447.

Kamienkowski, J. E., Carbajal, M. J., Bianchi, B., Sigman, M., & Shalom, D. E. (2018). Cumulative repetition effects across multiple readings of a word: evidence from eye movements. *Discourse Processes*, *55*(3), 256-271.

Kirkici, B., & Clahsen, H. (2013). Inflection and derivation in native and non-native language processing: Masked priming experiments on Turkish. *Bilingualism*, *16*(4), 776.

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 876.

Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and writing*, *12*(3), 191-218.

Marslen-Wilson, W. D., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes*, *23*(3), 394-421.

Münte, T. F., Say, T., Clahsen, H., Schiltz, K., & Kutas, M. (1999). Decomposition of morphologically complex words in English: Evidence from event-related brain potentials. *Cognitive Brain Research*, *7*(3), 241-253.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English?. *Reading research quarterly*, 304-330.

Neubauer, K. and Clahsen, H. (2009). Decomposition of inflected words in a second language. *Studies in Second Language Acquisition*, 31(3), pp.403-435.

Pacton, S., Foulin, J., Casalis, S., & Treiman, R., 2013. Children benefit from morphological relatedness when they learn to spell new words. *Frontiers in Psychology*, 4.

Pellicer-Sánchez, A. (2016). Incidental l2 acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition*, *38*(1), 97–130. https://doi.org/10.1017/S0272263115000224

Raveh, M. and Rueckl, J. (2000). Equivalent Effects of Inflected and Derived Primes: Long-Term Morphological Priming in Fragment Completion and Lexical Decision. *Journal of Memory and Language*, 42(1), 103-119.

Rodriguez-Fornells, A., Münte, T. and Clahsen, H. (2002). Morphological Priming in Spanish Verb Forms: An ERP Repetition Priming Study. *Journal of Cognitive Neuroscience*, 14(3), 443-454.

Royle, P., Drury, J. E., Bourguignon, N., & Steinhauer, K. (2012). The temporal dynamics of inflected word recognition: A masked ERP priming study of French verbs. *Neuropsychologia*, *50*(14), 3542-3553.

Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning*, *57*(2), 165–199. https://doi.org/10.1111/j.1467-9922.2007.00406.

Rott, S. (1999). The Effect Of Exposure Frequency On Intermediate Language Learners' incidental Vocabulary Acquisition And Retention Through Reading. *Studies In Second Language Acquisition*, *21*(4), 589-619.

Rueckl, J. & Aicher, K. (2008). Are CORNER and BROTHER morphologically complex? Not in the long term. *Language and Cognitive Processes*, 23(7-8), 972-1001.

Rueckl, J. G., & Raveh, M. (1999). The influence of morphological regularities on the dynamics of a connectionist network. *Brain and Language*, *68*(1-2), 110-117.

Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary learning and reading. *System*, *6*, 72–78. https://doi.org/10.1016/0346-251X(78)90027-1

Silva, R. And Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11(2), 245-260.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (Cambridge Applied Linguistics, pp. 3–32). Cambridge: Cambridge University Press. doi:10.1017/cbo9781139524780.003

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. *Morphological aspects of language processing*, *2*, 257-294.

Song, Mi-Jeong. (2007). Getting learners' attention: Typographical input enhancement, output, and their combined effects. *English Teaching, 62*(2), 193–215. https://doi.org/10.15858/engtea.62.2.200706.193

Stanners, R., Neiser, J., Hernon, W. and Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, 18(4), 399-412.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of verbal learning and verbal behavior*, *14*(6), 638-647.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader?.

Watanabe, Y. (1997). INPUT, INTAKE, AND RETENTION: Effects of Increased Processing on Incidental Learning of Foreign Language Vocabulary. *Studies in Second Language Acquisition*, *19*(3), 287–307. https://doi.org/10.1017/s027226319700301

Weyerts, H., Münte, T. F., Smid, H. G., & Heinze, H. J. (1996). Mental representations of morphologically complex words: an event-related potential study with adult humans. *Neuroscience Letters*, *206*(2-3), 125-128.

Winke, P.M., Godfroid, A., & Gass, S. M. (2013). Introduction to the special issue: Eye-movement recordings in second language research. *Studies in Second Language Acquisition, 35.* 205–212.

Wood, C. L., Mustian, A. L., & Cooke, N. L. (2012). Comparing whole-word and morphograph instruction during computer-assisted peer tutoring on students' acquisition and generalization of vocabulary. *Remedial and Special Education*, *33*(1), 39-47.

Wysocki, K., & Jenkins, J. R. (1987). Deriving word meanings through morphological generalization. *Reading Research Quarterly*, 66-81.