September 2019

Derivational Development: Derivational Word Processing in Three English-Speaking Populations

Lisa Suzanne Kemp
Louisiana State University and Agricultural and Mechanical College, lkemp4@lsu.edu

Recommended Citation

Kemp, Lisa Suzanne, "Derivational Development: Derivational Word Processing in Three English-Speaking Populations" (2019). *LSU Doctoral Dissertations*. 5040.
https://digitalcommons.lsu.edu/gradschool_dissertations/5040

DERIVATIONAL DEVELOPMENT: DERIVATIONAL WORD PROCESSING IN THREE ENGLISH-SPEAKING POPULATIONS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

by
Lisa Suzanne Kemp
B.A., James Madison University, 2006
M.A., California State University, Long Beach, 2013
December 2019

Acknowledgements

Thank you to the following people for their support during the preparation of this document. Dr. Janet L. McDonald, Dr. Emily Elliott, Dr. Neila Donovan, Dr. Christopher Cox, Dr. Eileen Haebig, Dr. Daniella Cash, Fatima Hernandez, Cameron Jones, Briana Sukert, Sanford Lemoine III, Mark & Beth Kemp and thank you to the Linguistics Faculty at CSULB for getting me here.

Table of Contents

List of Tables

List of Figures

Abstract

Native-English speaking adults use morphological decomposition to understand complex words (e.g. *farmer* becomes *farm-er*). Whether decomposition is driven by semantic organization is still unclear. It is also unclear whether ESL adults and elementary age children use the same word processing strategies as native speaking adults. This study tested an identical experimental procedure across three English-speaking populations: native speaking adults, non-native speaking adults and elementary age children. The first task tested how readers use base and suffix information in complex words and nonwords when the word featured only a base word, only a suffix, both a base and a suffix or neither. The second task was a masked priming task that evaluated how fast readers processed a word when paired with a transparent (farmer), opaque (corner) or simple (castle) word prime. For the first task, results showed both native and non-native adult English speakers use base and suffix information for English words and nonwords and vocabulary proficiency influences native speaker English word accuracy and non-native speaker nonword accuracy. Elementary age children did not use base or suffix information consistently for English words but used base word information for complex nonwords. Results for the second task, masked priming, showed a significant positive priming effect for transparent word pairs in all three language groups, significant positive priming in native speaking adults for opaque words and significant positive priming in children for simple words. Results for native speakers suggest a morphological decomposition strategy is obligatory. While non-native speakers of English use both base and suffix information when reading, morphological decomposition for this group is not obligatory. For the non-native speakers, age of acquisition did not interact with any of the experimental variables. Both tasks show elementary age children are still learning the rules of morphological decomposition and learning when morphological decomposition is an efficient strategy. The results of these studies have implications for vocabulary and literacy curriculum in both ESL and developing reader classrooms. Both learner groups would benefit from explicit suffix decomposition instruction as well as instruction regarding the semantic and grammatical role suffixes serve in word formation.

Chapter 1. Statement of Problem

This dissertation concerns the topic of visual word processing during reading, focusing on derivational word endings of complex words. After an introduction of the topic, I begin by reviewing the literature available on this topic for three populations: native English-speaking adults, non-native English-speaking adults, and native English-speaking elementary age children. I will then present the results of a study I conducted with native English-speaking undergraduates. This study replicates previous work on the topic and is the first experiment for my dissertation. The second and third experiments are conceptual replications of the native English-speaking group but with adult non-native speakers of English and elementary age, native English-speaking children. The goal of these experiments is to compare word processing strategies across groups and to examine how individual differences, specifically age of English acquisition and vocabulary proficiency, affect these strategies. The paper concludes with an in-depth comparison of the results from the three populations. I will address what similarities and differences exist among the groups across task and what these patterns mean in relation to complex word processing in English.

Complex words consist of small units called morphemes. In English, we create complex words by using three types of morphemes: simple words to create compound words (e.g. *paint* and *brush* make *paintbrush*), inflections (e.g. the *s* in *paints*), and derivations (e.g. the *er* in *painter*). Inflections are a marker for grammatical information in a sentence (e.g. past tense; noun-verb agreement), and derivations change the word meaning by changing its word class (e.g. *liquid* is a noun, and *liquify* is a verb). Inflectional forms of words are identical in meaning with added grammatical information (e.g. walk, walk*s*, walk*ed*) while derivational words can be related in meaning but are not identical. For example, the words *happy* and *happiness* are related but not identical in meaning.

How readers process compound words, inflections, and derivations may differ because each word formation type serves a different linguistic function. Compound words are formed to describe a new idea or entity. Inflections change the word's grammatical function but keep the same word meaning. Derivational word processing is challenging for the language learner because it asks them to learn how the suffix changes the base word meaning and word class (e.g. noun, verb, adjective). For example, a child must learn that adding *-ness* to the word *happy* changes the form (*happy* becomes *happi-ness*), the meaning (a happy feeling becomes an enduring emotional state), and the word class (from adjective to noun). Derivational word formation changes have the potential to be a source of difficulty in both first language and second language learning; for this reason, I chose to focus on this area of complex word processing for my dissertation.

The area I am specifically interested in is how a reader visually processes derived word forms. Readers use a variety of linguistic information, including spelling and sound structure, to process word meaning (Rastle, 2016). In complex word recognition, or words with multiple morphemes, an additional linguistic cue is word structure (e.g. a suffix or prefix). A core discussion in complex word recognition is even though a complex word like *charger* can be broken down into two parts (i.e charg-er), does the reader have to break the word down to understand it? Once the word is known, can the reader directly access the whole-word form from memory without breaking it down into separate parts?

In basic research, this question helps researchers examine how the reader uses linguistic information to access word knowledge. In applied research, this question helps educators understand what cognitive strategies are useful when learning new complex words and how these strategies change with expertise (e.g. literacy). In this paper, I identify whether derivational word processing strategies change with expertise within and across two age groups (children and adults) and within and across two language status groups (adult native speakers and adult non-native speakers).

Presently, researchers fall into one of two theoretical positions that attempt to explain how the reader uses and gives priority to linguistic information, such as spelling and semantics, during reading. Literature in this field defines whole-word access, also called lexical access, as when the reader retrieves a word without breaking it into parts. When the reader processes a word by its parts, the literature calls that morphological decomposition, or just decomposition. One position is that morphological decomposition has priority over whole-word access and that a complex word is always broken down by its morphemes before the meaning is accessed (Rastle & Davis, 2008; Rastle, 2016). In this model, a reader visually recognizes the word *walker*, breaks the word down into parts (walk-*er*), and then accesses the word meaning. This process occurs for any complex, multi-morphemic word.

The second position is that word meaning guides morphological decomposition and that a word is only broken down into parts if the suffix (or prefix) is connected to the word meaning (Giraudo & Grainger, 2000; Giraudo & Del Maso, 2016). In this model, the word meaning is mentally stored with all of its derivations underneath it. For example, *walk* would be stored as the head of the forms *walks*, *walked*, *walking*, and *walkers*. Once a word is read, word meaning is accessed, which leads to morphological decomposition if applicable. Derivational word processing studies look at three different types of words to test which model is correct.

Derivational suffix endings can relate to the base of the word. For example, the word *hunter* has two parts, *hunt* and *-er*, and it refers to a person who hunts. This is an example of a derived word form, or <u>transparent</u> word. However, a suffix does not have to refer to the base of the word. The word *corner* has a suffix, but it is not related to the word *corn*. In this sense it is not a true derived word. Even so, the reader can still break down the word further into two parts, a base and a suffix: *corn-er*. This kind of word is called a pseudoderived word, or <u>opaque</u> word.

A final category of words used in studies is called <u>form</u> words. These words are actually monomorphemic (they have one morpheme and no suffix) and cannot be broken down further into separate word parts. However, form words appear like they can be broken down because they contain the same spelling as a base word. An example of this word type is *castle*. It looks like it has two parts, *cast* and *-le*; but, the ending *-le* is not a suffix, so it is not a complex word. Form words are used as control words in derivational word processing research experiments.

I have described three word categories. Words that are derived from other simple words (*hunter*), words that have the form of a derived word but not the meaning relationship (*corner*), and words that look complex but have no form or meaning relationship (*castle*). I will use these three categories of words to talk about literature in the field as well as to describe the experimental items in my research proposal.

Methodology Background

Accuracy and response time data from lexical decision and priming repetition paradigms have shown that the adult reader processes words like *hunter*, *corner*, and *castle* differently and that these effects are replicable. One of the most common methods used for testing word processing strategies is repetition priming. In this technique, words and nonwords are shown one at a time on a computer screen while the participant completes a lexical decision task. Accuracy and response time measurements are recorded for each response as dependent measures.

The materials for priming experiments involve linguistically related words that are paired and inserted into the procedure. The manner in which these words are shown is explained after I first discuss the prime-target relationship. The first word in the pair is the prime and is designed to test whether the reader believes it is related in some way to the second word of the pair, the target. A fast response to the target would mean that the prime word made it easier for the reader to access the target word. This would mean that the reader has mentally organized the words in a way where the pair is linguistically related. An equivalent response time to the target compared to the rest of the trials would mean there is no linguistic relationship between the two words.

An example of a semantic prime pair would be *doctor-nurse*. In this case, reading *doctor* would make it faster for the participant to respond to the word *nurse* because they potentially share some of the same neural connections required for access. Since those connections were just activated after reading *doctor*, they are <u>primed</u> for activation of a following related word. An example of a derivational prime pair would be *hunter-hunt*. These two words are semantically related, so we can expect priming to occur for the target word *hunt*. This is the transparent pair relationship discussed above, in which the words are related in both form and meaning. However, what about for the pair *corner-corn*? These two words are not semantically related, but they do share the same base word. This is the opaque pair relationship. Priming studies have shown different effects for these two categories, depending on whether the prime word is masked or unmasked.

In masked priming, the prime word is shown on the computer screen very briefly, typically 50 milliseconds, which is too fast for the reader to consciously recall what was shown but slow enough to subconsciously read. Neutral visual stimuli is shown either before or after the prime word to clear the visual field, usually in the form of a row of hashmarks. This neutral stimulus is called a mask. A forward mask is a mask presented before the prime word, and a backward mask is a mask presented after the prime word. The paradigm used in most of the studies I review is that each trial has a forward mask shown for 500 milliseconds, a prime word shown for 50 milliseconds, and a target word shown for at least 1,500 milliseconds.

In unmasked priming, the prime word is shown at an equivalent duration as the other words in the procedure, and the word is consciously visible and readable. In this paradigm, a prime word is presented, and there may be filler words shown in between the prime and the target. How these paradigms identify whether whole-word access or morphological decomposition occurs during complex word processing is discussed in the literature review.

Chapter 2. Literature Review

The following two sections describe two theoretical positions that provide evidence for how readers process derivational suffix endings.

Obligatory Decomposition

Obligatory morphological decomposition refers to the theory that when someone reads a complex word, the reader breaks down the word by morphemes before sending that information to access word meaning. In other words, spelling information is processed before semantic information. According to this theory, this occurs for every multi-morphemic word regardless of its word meaning (Rastle, 2016).

Rastle, Davis, and New (2004) tested how native English-speaking undergraduates process transparent (*hunter*), opaque (*corner*), and form words (*castle*) in a masked priming task. Aside from the related prime-target pairs (hunter-HUNT), participants saw unrelated prime-target pairs (charger-HUNT) and unrelated prime-nonword pairs (charger-BLUP). Participants were asked to make a lexical decision as they viewed each target (Is this an English word, yes or no?). Primes were presented for 42 milliseconds, so the subjects were unaware of the primes in the experiment. To compute the priming effects for each transparent, opaque, and form category, the authors averaged the response times to targets for each condition and then subtracted the average response time for related primed targets from unrelated prime targets. They found a priming effect of 27 milliseconds on average for transparent prime-target pairs (hunter-HUNT), 22 milliseconds on average for opaque prime-target pairs (corner-CORN), and 4 milliseconds on average for form prime target pairs (castle-CAST). Both the transparent and opaque priming effect averages were significantly greater than those of the form priming effect.

The authors concluded that this provided support for the obligatory decomposition model because of the priming effect present in the opaque condition. If morphological decomposition only occurs if there is a semantic relationship, then there is no reason for *corner* to prime *corn* because the two words are not semantically related. However, if the reader mentally breaks down the morphemes in *corner*, then that would explain why participants are faster to respond to *corn* compared to the unrelated condition. Since the prime is only visible for 42 milliseconds, the priming effects show that morphological decomposition may occur early in the visual word recognition system.

Rastle and Davis (2008) conducted a review of masked priming studies that used the transparent, opaque, and form conditions between 1999 and 2008, and they found that seventeen of nineteen studies identified the presence of priming effects in the transparent and opaque conditions but not in the form condition. They also averaged the priming effects across all of the nineteen studies in each condition and got an average of 30 milliseconds for the transparent condition, 23 milliseconds in the opaque, and 2 milliseconds in the form condition. These results show there is repeated evidence for the obligatory morphological decomposition model using the masked priming technique.

In a unique study, Longtin and Meunier (2005) provide additional evidence for the obligatory decomposition model by testing whether a morphologically complex <u>nonword</u> prime would influence a real target word if the reader could interpret the nonword prime. The materials and procedure were conducted in French, but I am giving English translations in my examples. An example of a morphologically complex nonword is the word *quickify*. The base word is a valid word, and the suffix is a real suffix, but together they create a nonword. Since it is not a valid word, *quickify* should not prime *quick* if the visual word recognition system is semantically driven. To serve as controls, participants also saw related word prime-target pairs (quickly-QUICK) and unrelated prime-target pairs (other-QUICK). What they found was that the response time difference between the nonword prime (*quickify*) and the unrelated prime (41 milliseconds) was equal to the response time difference between the word related prime (*quickly*) and the unrelated prime (43 milliseconds).

The authors conclude that this result supports the obligatory decomposition model because it shows that the reader breaks down nonwords by morphemes to try and understand the word. If the evidence supported the whole-word access model, the nonword condition would not show a priming effect because the reader would break down the word into morphemes only after meaning was processed. Since the word is not a valid English word, meaning cannot be driving this priming effect.

More recent studies have attempted to tackle the whole-word versus morphological decomposition debate by isolating the time in which priming effects occur during processing. If semantics is driving processing, then semantic effects should occur first, and if morphological decomposition is driving processing, then these effects should occur first. Schmidtke, Matsuki, and Kuperman (2017) address this question in a series of lexical decision tasks using a “survival technique” analysis of traditional response times and eye tracking. The survival technique creates distributions for each variable level, and then levels are compared to mark the earliest time point in which an effect is taking place from the lexical decision response. This technique revealed which features of the word, word structure or meaning, show a response time effect first. The authors compared English transparent (*hunter*), opaque (*corner*), and form words (*castle*) and analyzed lexical decision response times. For the eye tracking data, they looked at first fixation duration times using only derived words (the transparent condition only) in order to accommodate the parameters of a larger study.

Variables relating to the structural properties of the word, such as stem and surface frequency, and variables relating to semantics, such as word and stem meaning, were coded for each word item. If form takes priority over semantics, then they should expect to see an effect in the survival distribution earliest for word structure and later in the time course for semantic features. For eye tracking, effects of fixation duration should be clear earlier in the distribution for surface frequency (structure) and later for word meaning (semantics).

The authors found that the effects for surface frequency emerged earlier in the lexical decision time course before word meaning for English transparent, opaque, and form words. In the eye tracking data, surface frequency effects in fixation duration for derived words occurred sooner (in milliseconds) than word meaning. This result provides evidence that word structure is processed before semantic information and that this process occurs for all derived word types, whether they relate to the base word (*hunter*) or not (*corner*).

In another study investigating the time course of word processing, Jared, Jouravlev, & Joanisse (2017) conducted a behavioral and ERP study comparing transparent, opaque, and form words using masked priming lexical and semantic decision tasks. In the lexical response data, they replicated the Rastle et al. (2004) pattern, in that transparent primes produced the largest priming effect, opaque primes showed a lesser priming effect, and form primes showed no effect. The difference between the priming effect for opaque and form primes was not significant for three of the four experiments in this study. Event-related potentials also showed priming effects for two time windows, 200-250 ms and 350-500 ms, in the transparent prime words but not the other two conditions.

For the semantic decision task, they asked participants to identify if a word on a screen was a sea creature, yes or no. ERP data was also recorded. The word materials were the same from the masked priming task, but they added a manipulation where some of the words had the morpheme bolded accurately (**book**ish) and some had them bolded inaccurately (boo**kish**). If morphological units are processed first, this should be shown in the form of slow response times for inaccurately bolded words. They found this was the case but only for transparent words and not opaque or form words.

Here, the lexical decision data provides support for the obligatory decomposition model, but the ERP and semantic decision data does not. In the latter two, only the semantically derived words (*hunter*) are priming the target. It is possible that the cumulative RT effect of the lexical response includes additional processing not measured in the ERP data. Further ERP research is necessary to break down what processing events are included in the lexical decision response times and the order in which they occur.

One area that is lacking in the literature is a discussion of how individual differences, including language proficiency, influence morphological decomposition strategies. A few studies have tackled this question, and it is discussed after the following section.

Non-Obligatory Decomposition

The studies discussed below propose an explanation for masked priming effects that do not support an obligatory decomposition model. I do not call this section a whole-word processing or semantically driven model because while these studies do provide evidence against the obligatory decomposition model, they do not entirely support the whole-word access model. Marslen-Wilson, Tyler, Waksler, and Older (1994) used cross-modal repetition priming to investigate the degree to which semantic overlap between the prime and target pair is necessary to reveal a priming effect. In cross-modal priming, the prime word is heard through headphones, and the target is viewed on-screen immediately after the prime is heard. The authors chose this priming method in order to account for any priming effects that might be exclusive to the visual system. The article includes several experiments that manipulate semantic, morphological, and phonetic variables within the prime-target pair. However, I will only review the section that compared the variables of interest to this research.

In one experiment, they asked whether a transparent prime-target pair (hunter-HUNT) that shared a semantic and morphological relationship would show equivalent priming effects to an opaque prime-target pair (corner-CORN) that only had the morphological relationship. They did not use a control prime-target pair (castle-CAST). They found the difference between the unrelated and related prime-target pairs in the transparent condition was significantly different from zero at 35 milliseconds; however, the difference in the opaque condition was not significantly different from zero at 15 milliseconds. They replicated this lack of priming effect in a follow-up experiment that featured more trials, and they found no priming effect between the unrelated and related primes in the opaque condition (corner-CORN). They concluded that semantic relatedness between the prime and target pair (hunter-HUNT) is necessary for a priming effect to occur and that morphological relatedness without semantic relatedness is not enough to produce a priming effect. In this case, morphological decomposition only occurs when there is a semantic relationship. It is also possible that the target is primed solely because of the semantic relationship and not because of the morphological relationship.

Rastle's response to this article is that whenever the participant is aware of the prime, whether using unmasked priming or cross-modal priming, the mechanism used to process the word is semantically driven and is occurring after the morphological processing window. She argues that unmasked priming is necessary to reveal the early orthographic processing stage in these experiments.

Diependaele, Sandra, and Grainger (2005) is a study that stands out among the masked priming studies that have consistently found priming effects in the opaque prime-target condition (corner-corn) that were reviewed in Rastle and Davis (2008). The authors created an elaborate, experimental procedure that counterbalanced masked priming and cross-modal priming trials, comparing the transparent, opaque, and form conditions in French derived word forms. In addition, the presentation of the prime duration was also counterbalanced. The duration of the prime was manipulated so that one third of the primes were presented for 13 milliseconds, another third for 40 milliseconds, and the final third were presented for 67 milliseconds. The final procedure involved a 2 Modality (visual/auditory) X 3 Prime (transparent/opaque/form) X 2 Relatedness (unrelated/related prime) X 3 Prime Duration (13/40/67 ms) repeated-measures design. With this set up, the researchers were able to test the degree to which priming effects present in the transparent, opaque, and form conditions were specific to any one modality or prime duration.

For both masked priming trials and cross-modal trials, they found a significant priming effect for all three prime-target pairs—transparent (hunter-hunt), opaque (corner-corn), and form (castle-cast)—but only when the prime duration lasted for 67 milliseconds. In the masked priming trials, the 40-millisecond prime duration produced a priming effect only in the transparent condition, not in the opaque or form conditions. This contradicts what Rastle and Davis (2008) claim is special about the masked priming paradigm because it is within the 40 to 50-millisecond priming window in which morphological decomposition is supposed to occur during visual word processing. The 40-millisecond prime duration on the cross-modal trials had a negative priming effect in the conditions that was not statistically significant. The 13-millisecond window most likely showed no priming effects in either modality because it was too brief to be processed.

What is interesting about this study is that it does not support the obligatory decomposition view, but it also does not support the view that semantics must drive lexical decomposition. Obligatory decomposition is not occurring in the transparent and opaque conditions in the 40-millisecond prime window as it should, but it did occur in the 67-millisecond masked priming and cross-modal priming window. The whole-word access model (semantics drives decomposition) states that priming in the opaque condition is not possible. The authors conclude that there may be two word recognition systems at work that allow for both phenomena to occur. One system engages obligatory decomposition (the 67-millisecond window), and one engages in semantically driven priming (the transparent condition in the 40-millisecond window). However, this study raises questions: If obligatory decomposition occurs early in the word recognition system (Rastle, 2016), then why does it occur in a later time window in this experiment?

Feldman, Milin, Cho, del Prado Martin, and O'Connor (2015) offers more evidence in regard to the question of prime duration by manipulating the stimulus onset asynchrony, or SOA, of the masked word prime in the transparent, opaque, and form prime-target pair conditions from 34 to 100 milliseconds. If their results support the obligatory decomposition view, then priming effects should occur in both the transparent and opaque conditions in the early SOA times. If the results support a semantics driven decomposition view, then priming effects should only be present in the transparent prime-target condition across all SOAs.

In the analysis, they calculated the response time effects a little differently than other studies. They held the target word constant between transparent and opaque conditions, and they subtracted the response time of the opaque condition (sneaker-SNEAK) from the transparent condition (sneaky-SNEAK) at each SOA (34, 48, 67, 84 and 100 ms) instead of subtracting the related (hunter-HUNT) from the unrelated condition (charger-HUNT). They found that the first and last SOA produced equivalent response times between the opaque and transparent conditions. The 48 and 67-millisecond SOA showed a 30-millisecond difference between the opaque and transparent conditions with response times fastest in the transparent condition. The 87-millisecond SOA revealed a 12-milisecond difference between the transparent and opaque conditions, with the transparent condition faster than the opaque condition.

In four of the five SOAs, the opaque condition (corner-CORN) had faster response times on average than the form condition (castle-CAST). This is important since, according to the whole-word access model, opaque priming effects should be equivalent to the form condition because semantics drives decomposition; and, if there is not a semantic relationship, then decomposition will not occur. However, Rastle claims that the opaque condition does not have to be identical to the transparent condition for it to be evidence of morphological decomposition. She instead argues the priming effect between unrelated and related prime types in the opaque condition has to be different than zero for it to be considered evidence (Rastle & Davis, 2008). It is concerning that in the Feldman et al. (2015) study, the opaque condition shows no response time difference compared to the form condition in the 48-millisecond SOA. It is this SOA duration where so many studies identified in Rastle and Davis (2008) provide an opaque priming effect.

The Feldman et al. (2015) and Diependaele et al. (2005) studies have similar results in that the transparent condition had priming effects at the 48-millisecond window, and both the transparent and opaque conditions had priming effects at the 67-millisecond window. The Feldman et al. (2015) study, in my view, provides evidence for a dual system mechanism discussed in Diependaele et al. (2005), in which both obligatory decomposition and semantically driven decomposition are possible.

The studies reviewed in this section highlight the difficulty of identifying consistent results in the opaque prime-target pair condition. By manipulating priming techniques and prime durations, the last two studies are able to show that both semantic and morphological information present in a complex word help the reader process the word. However, it is not possible to conclude that obligatory decomposition is an accurate representation of complex word processing. It is likely that complex word processing strategies depend on the materials, methods, and population used in the experiment.

The evidence supporting the obligatory decomposition model is clearly replicable but is relegated to a specific priming paradigm. It is not obvious whether morphological decomposition must occur in order for meaning to be accessed. If that were the case, then priming effects should be present in the opaque condition in experiments that do not use a masked priming paradigm. It is clear there is something special about the transparent word condition in which the combination of semantic and morphological relatedness creates a unique connection to the target word. It is this condition that has the largest priming effect in the studies reviewed in Rastle and Davis (2008), and it is the condition that shows priming effects in paradigms outside the masked priming paradigm (Jared et al., 2017). It does not mean that morphological decomposition is semantically driven, but it does mean that visual word processing may be more complex than either theoretical model suggests.

Individual Differences

Before discussing categorical populations that might differ in complex word processing, I would like to review some studies that have looked at how individual differences with native adult speakers influence processing. Researchers have begun to explore how individual differences, specifically proficiency, might influence word processing strategies.

Andrews & Lo (2013) asked whether a participant’s reading or spelling proficiency influenced the likelihood of finding a priming effect in a masked priming task. The participants consisted of native English-speaking undergraduate students, and their individual differences were assessed using a vocabulary test and a verbal and written spelling test. The authors used the same derivational variables as the studies above and came up with transparent, opaque, and form prime-target pairs for the materials. The authors found the same priming effect as previous studies, with transparent primes producing the largest priming effect, opaque in the middle, and form showing no priming effect. Like previous work, opaque primes produced a higher priming effect than the form primes on average, but the effect was not statistically significant in comparison to either the transparent or form condition.

For the individual difference measures, the authors combined the performances on all three proficiency tests for each participant and compared it to the priming effects. There was no interaction between proficiency and priming for that comparison. Next, the authors compared the participant’s performance on vocabulary versus spelling and plotted the participant’s relationship between the two measures. A high score on the plot meant that the participant scored better at spelling than vocabulary, and a low score meant that vocabulary was better than spelling. They found that those with better spelling scores compared to vocabulary displayed a significantly larger priming effect on opaque words but not transparent words. However, if a participant was better at vocabulary than spelling, they displayed a significantly larger priming effect for transparent words but not opaque words.

This relationship is interesting because it ties to the linguistic properties of the two variable conditions. Transparent primes share both a semantic and a form relationship to the target whereas opaque primes share a false semantic connection but also a related form to the target word. Since spelling skills are based on orthography, it makes sense that good spellers would display an advantage for opaque related primes. Likewise, since vocabulary skills are related to semantic access, it also makes sense that participants with higher vocabulary performance than spelling performance display a priming advantage for transparent primes. These results show that derivational word processing effects vary and change based on the skill set and background of the participant population.

Medeiros & Duñabeitia (2016) evaluated individual differences in Spanish-speaking participants who completed a masked priming lexical decision task with Spanish derived complex words. They also completed a lexical decision with simple words to evaluate individual response time averages. The primes in this experiment were matched to targets by the amount of spelling overlap: A related prime shared the same suffix as the target (e.g. charger-hunter), and an unrelated prime did not share the same ending. Participants were split into two groups, fast readers and slow readers, based on the median response time performance for that task. They found that fast readers did not show a significant priming effect (unrelated-related word primes) for any condition, but the slow readers did.

This study shows that participants who take more time to process the words may influence the effects shown in the cumulative response times from the lexical decision task. It is possible that participants who read quickly are responding before any word processing differences are reflected in the lexical decision response times. In other words, we need to make the task harder for the fast readers in order to reveal processing differences between the conditions.

The studies referenced in this section reveal that researchers are thinking about the relationship between experience and complex word processing. However, more replication in this area is needed to understand the degree to which proficiency affects derivation word processing. I am specifically interested in the factor of age of English acquisition and how it relates to the use of morphological decomposition.

The following two sections discuss how derivational word processing is reflected in non-native adult speakers and native English-speaking children.

Second Language Acquisition

Cognitive scientists know the brain changes with experience and age, so it is possible that the way readers process complex words is not consistent or constant. Variables like vocabulary size, age, and literacy status could influence whether whole-word processing or morphological decomposition is helpful when reading. In second language learning, the morphological complexity of a native speaker's first language may influence how they process complex English words. For example, Mandarin has many compound words but not many suffixes (Wang & Verhoeven, 2015). This difference might mean that a native Chinese speaker is more likely to engage in whole-word access since morphological decomposition is uncommon in their native language. Non-native speakers (L2 speakers) learn English later in life, which influences their level of mastery, and native and non-native speakers' respective levels of English proficiency may influence their complex word processing. Both native and non-native speakers' language experience background may also influence complex word processing. The following studies introduce how researchers have examined the whole-word access and morphological decomposition question in non-native speakers.

Diependaele, Duñabeitia, Morris, and Keuleers (2011) challenge the idea that L2 speakers of English process complex words using a whole-word strategy. As a baseline, native English-speaking participants completed a masked priming task using transparent, opaque, and form word pairs that feature both unrelated and related primes. Using a linear effects model, the results showed a similar descending pattern of priming effects to Rastle and Davis (2008): Transparent word pairs facilitated an average reaction time of 36 milliseconds, opaque word pairs took 15 milliseconds on average, and form word pairs saw no facilitation with a 1-millisecond difference between unrelated and related primes for the same target word. This data supports the priming effect present in most English L1 priming studies (Rastle & Davis, 2008). Next, they conducted two studies of masked priming effects on adult L2 learners of English, one with Spanish-English bilinguals and the second with Dutch-English bilinguals. Native Spanish speakers had an average age of English acquisition at seven years and completed the same experiment as the native English speakers. They found the L2 speakers matched controls for the transparent priming effect at 35 milliseconds on average as well as a smaller priming effect of 25 milliseconds for the opaque word pairs. For the form pairs (castle-CAST), the L2 speakers did show significant priming between unrelated and related primes with a 14-millisecond advantage in the related condition on average.

Next, Dutch-English bilinguals with an average age of English acquisition at 11 years also completed the experiment. This group showed an almost identical priming pattern to the Spanish speakers with a 35-millisecond facilitation for the transparent condition, a 26-millisecond facilitation for the opaque condition, and a 14-millisecond facilitation for the form word pairs. These results show that, like native speakers, the English L2 speakers are sensitive to the difference between transparent and opaque primes; but, unlike native speakers, they are also sensitive to form word primes. This shows that non-native speakers engage in a morphological decomposition strategy that is not semantically driven since neither opaque nor form primes have semantic relationships to their targets.

In another language group, Li, Taft, & Xu (2017) found similar results as Diependaele et al. (2011), recruiting Mandarin-English participants to complete a masked priming task featuring the same transparent, opaque, and form word pair conditions. The authors suggested that Mandarin-English bilinguals might engage in less English lexical decomposition since Mandarin characters feature very few derivational morphemes. If the L2 group shows no priming effect differences between the three word conditions, we might assume that Mandarin-English bilinguals process English derived words using a whole-word strategy to match how they process Mandarin. However, if the participant group processes the words similarly to native English speakers, we might expect a facilitation effect in the transparent and opaque conditions but not the form condition. The authors of the study were also interested in word processing as a function of English language proficiency, so participants were separated into a high or low proficiency groups before the study, based on percentage of daily English use and accuracy on an English vocabulary test.

Response time averages showed that in the low proficiency group, transparent (hunter-HUNT) and form (castle-CAST) related primes produced an equivalent priming effect at 57 and 54 milliseconds respectively. The opaque related prime (corner-CORN) showed a lower priming effect at 31 milliseconds. In the high proficiency group, the RT pattern is more native-like, with the transparent condition showing a priming effect of 59 milliseconds and with the opaque and form primes producing a priming effect of 26 and 22 milliseconds respectively.

It is in the form condition where the differences in proficiency are clear in respect to the degree of morphological decomposition. The low proficiency participants engaged in morphological decomposition for the form primes even though there are no real English suffixes in this condition (e.g. castle) and no decomposition is required. In comparison, the high proficiency speakers reflected the participants in Diependaele et al. (2011), where the transparent and opaque conditions looked native-like in their priming effects, but the form condition is still being primed.

These two studies show that non-native speakers may be overusing morphological decomposition as a strategy to understand complex English words. Participants from the Diependaele et al. (2011) study acquired English proficiency in adolescence but started learning the language in elementary school. Both participant groups from Li et al. (2017) began learning English at grade 4. I am interested in whether non-native speakers of English who are highly proficient and have an earlier age of acquisition (e.g. age 5) exhibit the same overuse of morphological decomposition. Or is it a strategy that is abandoned once a certain level of mastery is reached?

In a study that used a different paradigm from masked priming—a non-priming lexical decision task—Casalis, Commissaire, and Duncan (2015) asked the degree to which non-native speakers of English are sensitive to complex word structure. French-English bilinguals were identified as low or high proficiency speakers before the experiment, based on study abroad experience and accuracy on a word translation task. For the task, English words and pseudowords were chosen for four different conditions that varied by their presence or absence of a real English word or real English suffix. W+S+ English words had either a base word and a suffix (e.g. bony), W+S- words had a base word and no suffix (e.g. bullet), W-S+ words had no base word and a suffix (e.g. hoover), and W-S- words had no base word and no suffix (e.g. elbow). Pseudowords were developed and assigned to the same four conditions.

The hypothesis for the study was that if L2 speakers do not engage in morphological decomposition as suggested by Silva and Clahsen (2008), then participants should display no differences in response times between the four conditions. However, the authors found that participants were more accurate for English words that featured a suffix compared to those that did not (e.g. clobber vs. cobble). Low proficiency speakers were more accurate and faster to respond to English words that featured a base word (e.g. catch vs. watch), but items with base words did not help high proficiency speakers since their accuracy was already high.

For nonwords, high proficiency speakers were better at rejecting nonwords than the low proficiency group. If a nonword featured a real base word (e.g. cabol) or real suffix (e.g. clockage), this was more likely to contribute to an incorrect selection of the word as a real English word for both proficiency groups.

Overall, suffixes applied to English words helped readers make a correct selection, and suffixes applied to nonwords hurt their chances of making a correct rejection. This shows that the participants were sensitive to word structure during the task and, therefore, could not have processed the words using a whole-word access method.

The results show that both proficiency groups are sensitive to the internal structure of the English words and nonwords. English suffixes seem to help word recognition, which suggests that the participants engaged in decomposition. Words with base words helped the low proficiency speakers recognize the word, which means that as English vocabulary knowledge increases, base word recognition may be less important. It is also possible that the word items were not difficult enough for the high proficiency participants to reveal any base word effects. The accuracy data from the nonwords shows that morphological decomposition is an important part of L2 English word reading since both proficiency groups were sensitive to nonwords that featured real base words and/or real English suffixes. Information from this study cannot provide evidence for the obligatory decomposition model because even though the native speakers are using a morphological decomposition strategy, we do not know if it is used in combination with any other kind of word processing, including semantic or phonological.

The studies reviewed in this section show that non-native speakers of English engage in both semantic and morphological processing of complex words. We can say that non-native speakers do not engage in a whole-word-access-only approach, in which semantics must drive decomposition. If that were the case, then non-native speakers should not show priming effects in the form condition of priming tasks (castle-CAST), and they should also not be sensitive to the internal word structure of nonwords (e.g. clockage). The studies reviewed give a little information on how morphological decomposition changes with proficiency.

In Casalis et al. (2015b), low proficiency participants used both base words and suffix structures to make correct decisions about English words, but high proficiency participants only benefitted from the suffix. This suggests that low proficiency readers may overuse morphological decomposition as a strategy for complex word processing. Another significant factor that needs to be explored is age of acquisition. The age at which a non-native speaker acquires English has not yet been explored in regard to how it influences complex word processing. Comparing groups based on age of acquisition could offer information about how morphological decomposition strategies are used once English mastery has been reached in a non-native speaker.

First Language Acquisition

Derivational word acquisition in young readers is key for literacy achievement. Children must understand the rules of when a suffix is appropriate for a word ending and how the meaning changes once the suffix is applied to a base word (Carlisle, 2000). A handful of studies have experimentally tested how children process complex derived word forms. The studies reviewed below address the same questions as with the adult population but with a specific focus on how these relationships change as children become more skilled in reading. Children might differ from adults in how they process complex word forms because, like non-native speakers, they are still learning the rules of English. They may overuse the morphological decomposition strategy like non-native speakers, or they might use a whole-word strategy because they are unable to identify the word structure in derived word forms.

In elementary age child readers, Carlisle and Stone (2005) found that third and fourth graders were significantly faster at reading transparent words, like *hunter*, compared to opaque words, like *corner*, whereas fifth and sixth graders read these word types at the same speed. This means that less literate children process transparent and opaque words differently than more literate children. Whether this means they are using a whole-word or morphological decomposition strategy is unclear.

In order to tackle the decomposition question in younger age groups, Mahony, Singson, & Mann (2000) developed an easy task that asks the student whether a word pair is related or not. In their study, they tested grades three through six, and 20 of the word pairs were related while 20 word pairs were not. They completed the test orally and in written form. Many of the variables in the word materials were designed to evaluate recognition of sound changes in the pair, like *numb-numbers*, which is a topic outside the scope of the study. However, one relevant finding for this paper is that performance on the test improved with age, as third graders were more likely to incorrectly accept opaque word pairs, like *bug-buggy*, or words that share a morphemic but not semantic relationship. After fourth grade, readers are able to correctly understand that these two words are not related in meaning. These results show that with increased reading experience, morphological decomposition is no longer used to identify whether two words are related in an explicit task. This does not indicate whether *buggy* would prime *bug* in a masked priming task.

The following two studies introduce an important methodology that I will use in each of my three experiments. The first study, unfortunately, lacked power and failed to recruit enough participants. However, the second study was able to conduct a sufficient between-subjects comparison.

Gray, Quémart, and Casalis (2011) compared 20 above average readers and ten below average readers in the third or fourth grade on their ability to complete a lexical decision task that asked them to quickly decide if a word on a screen was an English word. These conditions were identical to Casalis et al. (2015a), reviewed in the second language acquisition section. W+S+ words had a base word and a suffix (e.g. farmer), W+S- words had a base word and no suffix (e.g. window), W-S+ words had no base word and a suffix (e.g. murder), and W-S- words had no base word and no suffix (e.g. narrow). The nonwords for the experiment included the same four categories. Individual difference measures included fluid intelligence (Raven's Colored Progressive Matrices), and a word reading test and an expressive vocabulary test were also completed.

Results showed that above average readers did not make more accurate lexical decisions for words with a real base word when there was a suffix (*hunter*) compared to when there was no suffix (*window*). When there was no base word, this group was more accurate for words with a suffix (*murder*) compared to words without a suffix (*narrow*). For nonwords, the good readers significantly failed to reject nonwords with a real base word and a real suffix more than the other three conditions. In other words, accuracy was lowest for nonwords that seemed the most English-like.

This indicates that the children are sensitive to word structure for both English and nonwords. These results match the high proficiency adult speakers in Casalis et al. (2015a) in that the high proficiency adults only used the suffix to help make the lexical decision for English words, but both base word and suffix structure hurt their accuracy for nonwords.

Casalis, Quémart, & Duncan (2015) conducted an identical lexical decision task with English (British) and French fourth graders and were able to recruit more participants for their study (n=40 and n=32 respectively). I will only report on the English group for this section. The procedure and conditions were identical to the previous study. In this group, the accuracy results matched the previous study in that the fourth graders were most accurate on words with no base word but with a real suffix (*murder*). They also were slowest to correctly accept words that had a real base word, either with or without a suffix. Also, like the previous study, the students had the most difficulty rejecting nonwords that were the most English-like, those with a real word and a real suffix.

These results suggest older children use morphological decomposition as a strategy to understand English words and they also use this same strategy for nonwords that can be morphologically decomposed. It is possible that the third graders have not yet mastered English suffixes to use them as a strategy for word processing. Like the lexical decision task in the non-native group, these results cannot explain whether morphological decomposition is obligatory, but it can indicate whether decomposition is a strategy that is used to understand words during the learning process and how this strategy changes with experience.

One study addresses this question directly and tests how different age groups perform on the same lexical decision study. Dawson, Rastle, and Ricketts (2018) conducted a cross-sectional lexical decision study investigating nonword decomposition. They recruited four age groups: elementary age children, younger adolescents, older adolescents, and adults. The authors also asked the participants to complete three individual difference measures, including a nonverbal reasoning test, an oral vocabulary test, and a reading test. These measures were used for descriptive purposes and were not included in the statistical model during analysis.

The purpose of the study was to investigate how the participants processed two types of nonwords. The first type of nonword featured a valid base word with a valid suffix that created a nonword (e.g. earist). This condition is the same as the W+S+ (base present, suffix present) condition in Casalis et al. (2015b). The second nonword type featured a valid word with a made-up word ending to create a nonword (e.g. earilt). This condition is the same as the W+S- (base present, suffix absent) condition in Casalis et al. (2015b). The authors wanted to know if the age groups would process the W+S+ faster because decomposition was facilitating the lexical decision both in accuracy and response time.

The accuracy results showed that correct rejections increased as age increased; however, there was a significant difference in accuracy between the younger and older age groups. The elementary age children and younger adolescents did not differ from one another, but they were significantly less accurate than the older adolescents and adults in the W+S+ condition. Likewise, the older adolescents and adults did not differ from one another.

Response time data showed that older adolescents responded significantly faster than younger adolescents, and adults responded significantly faster than older adolescents. There was, however, no difference in response times between younger adolescents and children. Since response times are evidence of lexical decomposition, the authors conclude that the W+S+ condition is only facilitating responses in the older adolescent and adult group. They suggest this may occur because the younger age groups are still learning how suffixes generalize to other words, so decomposition is not practiced at this age. Their explanation for why the W+S+ nonwords facilitated lexical decision in the older adolescents is that readers of this age repeatedly encounter more complex words in textbooks and advanced novels, so decomposition has become a more useful cognitive strategy.

Lexical decision tasks can only state whether lexical decomposition has occurred, and this paradigm cannot isolate the time point of that process or whether it is driven by semantics. However, this study has shown that lexical decomposition is a skill that is built with increased exposure to complex words and that is useful during visual word recognition.

Beyersmann, Castles, and Coltheart (2012) gives further evidence that lexical decomposition is a strategy that develops with age. The authors used the masked priming paradigm to test transparent, opaque, and form prime-target pairs in an adult participant group and in an elementary age group with children in grades three and five. The prime duration was 50 milliseconds. The authors found the same priming effect pattern as Rastle et al. (2004) in the adults but only found evidence for transparent priming in both elementary grade levels. In fact, at grade three the opaque condition showed negative priming effects, and at grade five both the opaque and form (castle-cast) conditions showed negative priming effects. These inconsistent patterns show that children are still mastering the rules of derivational suffixes.

It is difficult to identify consistent priming effects in young children on a masked priming task, but Carlisle and Stone (2005) used a simpler implicit task to examine derivational word processing in children. In this study, the authors measured how fast and how accurately elementary students could read a list of derived words that were manipulated based on semantic transparency and word frequency. Words were presented on a screen, and response latency was recorded by the experimental software. For the materials, the authors applied the transparency manipulation based on whether the word is transparent (e.g. shady) or opaque (e.g. lady). Students were grouped into lower (grades 2 and 3) and upper (grades 5 and 6) elementary groups. A reaction time task was also given as an individual difference measure, and students were screened on a reading and vocabulary measure before participating. The upper students were better at the reaction time task, but it did not interact or mediate the relationship with their word reading performance.

The results showed that the upper levels were faster and more accurate at the reading task than the lower levels. Both grade groups were more accurate on the transparent words compared to the opaque words. They found the lower grade level was significantly faster at reading the transparent words compared to opaque words, but the upper grade read both types of words at the same speed.

The individual difference correlations showed that reading performance on a letter-word identification and word attack task significantly correlated with the lower grades' speed and accuracy performance on all of the word forms. Reading performance correlated only with low frequency accuracy scores for the upper grade level.

These results are contradictory to the lexical decision results in Casalis et al. (2015b), wherein the younger students with less reading experience were not sensitive to the word differences in the lexical decision tasks. However, in this study, the younger students are sensitive to the differences of the transparent (*hunter*) and opaque (*corner*) words, but the older students are not. However, in the Carlisle and Stone (2005) study, the two-word types differ by semantics, not structure. It is possible this task identifies that the younger students process *hunter* and *corner* differently because they are using a semantic, whole-word approach whereas the older students process them the same because they are using a morphological decomposition approach. Future research should tease apart how proficiency in children contributes to processing differences in words that vary by structure (Casalis et al., 2015b) and in words that vary by semantic relationship rather than structure (Carlisle and Stone, 2005).

Overall, the evidence from these experiments indicates that before high school, students may use semantically driven morphological decomposition, or whole-word access. This is shown in Carlisle and Stone (2005), where the younger students had significantly different reading times for transparent and opaque words while the older students did not. Also, in the cross-sectional studies, the young age groups did not morphologically decompose conditions that featured only word structure similarities and no semantic similarities, but the adults did

(Beyersmann et al., 2012; Dawson et al., 2018). However, if the method only compares structural differences, as in the W+/- and S+/- manipulations in Duncan et al. (2011) and Casalis et al. (2015b), the evidence shows that children are able to use morphological decomposition as a tool for word processing. There is not enough evidence to show how proficiency changes morphological decomposition patterns, but results from Carlisle and Stone (2005) suggest that whole-word access decreases with reading experience.

Chapter 3. Experiment 1. Adult Native Speakers

The goal of this experiment was to provide a baseline with which to compare later experiments by testing adult native-English speakers. This group completed a simple lexical decision task and a masked priming task. The lexical decision task adapted the W+/- and S+/- conditions from Casalis et al. (2015a) and the masked priming task featured the transparent, opaque and form conditions also discussed in the literature review.

My research questions for experiment #1 were whether adult native English speakers would be sensitive to English word and nonword structure changes in the lexical decision task as revealed by accuracy and response time measurements and whether they would show response time positive priming effects in the opaque condition (corner-CORN) on the priming task. If they did show positive priming effects in this condition, this would suggest morphological decomposition is not semantically driven in this population. Following previous literature and the studies from which these materials were adapted, I hypothesized that this group would be sensitive to word structure changes in the lexical decision task and that they would produce priming effects in the opaque condition. These results would show adult native speakers of English mentally decompose English morphemes and that this decomposition is not rigidly tied to word meaning.

Native-English speaking undergraduates were recruited from the LSU psychology research pool. In addition to completing a language background questionnaire, participants completed an English vocabulary test and a Reading Span test to measure individual differences, as well as a lexical decision task and a masked priming task to measure differences in lexical access of derivational words.

Method

Participants

Seventy-eight native English-speaking undergraduate students at a university in the Southern United States completed the tasks. All measures were conducted in English using English language materials. Participant ages ranged from 17 to 27 years (*M*=19.3, S*D*=1.73) and consisted of 61 females and 17 males. Seventy-two participants reported studying at least one foreign language in a classroom setting either in high school or college and all participants reported speaking English at least 93% of the time during the day on average. Reading habits and language background information was also gathered.

Materials and Procedure

All tasks except the reading span task were completed on a Mac computer using SuperLab software **(**Haxby, Parasuraman, Lalonde, & Abboud, 1993). The Reading Span task was completed on a Microsoft computer in the same lab room using E-Prime (Psychology Software Tools, Inc., 2016). All tasks were completed in a single day within a half-hour time frame.

English Vocabulary Test

Subjects completed a difficult multiple-choice English vocabulary test with 40-word items on a white screen with black text. Subjects were asked to look at an English word and select the best definition out of five options. Participants had as much time as they needed to select the correct answer. To choose their answer, participants used the number keys 1-5 to represent answer options A-E on the screen. Words were selected based on study guides available for the GRE and the complete materials for the test are included in Appendix D.

Reading Span Test

Subjects completed the shortened version of the reading span test from the Engle Lab (Oswald, McAbee, Redick, & Hambrick, 2015). This assessment is a dual task memory test that requires participants to read an English sentence and select whether the sentence is grammatically correct, yes or no. Participants are then shown a letter which they are asked to remember. This procedure continues until the participant is shown a screen with multiple letter options and they are required to select the letters they saw in the correct order they saw them. To earn a point for the letter, participants must select the correct letter recalled in the correct order they saw the letter. They also must earn a score of 80% or higher on the grammaticality judgment portion for their memory scores to count. The number of sentence/letter combination the participant is asked to complete varies depending on the trial. For this experiment, I used the partial span score which is all correct letters in the correct order counted, even if not all letters in the trial were correct. This task was given last in the experimental procedure and not all participants were able to complete this task for time reasons. For example, if a participant arrived late or if they took a long time to complete the other three tasks, the reading span task was not administered.

Lexical Decision Task

The words for this task were adapted from Casalis, Commissaire and Duncan (2015a) and are listed in Appendix E. There are two variables (base word and suffix) and two levels in each variable (present + or absent -). Casalis et al. (2015a) created four English word lists to match the features of each level. The first list, W+S+, contains words that have an English base word and an English suffix (e.g. cloud-y). The second list contains words that have an English base word and a letter string that is not a suffix, W+S- (e.g. jack-et). The third list has no base word but contains a suffix, W-S+ (e.g. dut-y) and the last list has neither an English word or a suffix, W-S- (e.g. elbow). They also created a list of nonwords to match each English word list and were assigned the same list labels: W+S+ (e.g. farl-y), W+S- (e.g. doll-et), W-S+ (e.g. dral-y) and W-S- (e.g. forrow).

Each English and nonword list had twenty words for a total of eighty items. All words and nonwords were matched for length and number of phonemes or word sounds in a previous article and the materials used were directly adapted from this article (Casalis, Commissaire & Duncan, 2015b). Five words were replaced in the W-S+ English condition because the Casalis et al. (2015b) article marked them as assigned to the wrong condition. The words belly, easter, luggage, marry and pepper were replaced with the words mirage, docile, lobster, future and leather. Participants responded to twelve practice words before beginning the experimental portion and the experimental words and nonwords were randomly presented to each participant following the practice session.

For the lexical decision task, words were displayed in black text on a white background in the middle of the screen. Words were shown in courier font and in text size 24. Participants had 5000 milliseconds to respond to the string on the screen before the next item was shown. There was a blank screen in between each word item that lasted for 1000 milliseconds. If the word left the screen before a response was entered, it was marked as wrong. This was not a significant problem during testing and occurred on a total of three trials for both English words and nonwords. Participants received the same instructions as the Casalis et al. (2015a) article in order to prevent long response times. Specifically, subjects were told all the English words were common words and that they were to enter their response on the keyboard as soon as they recognized the string as a word or a nonword by pushing the “G” key for yes and the “H” key for no. The “G” key had a sticker on the key with a Y on it and the “H” key with an N on it.

Masked Priming Task

Materials for the masked priming task were adapted from Li, Taft and Xu (2017) and are included in Appendix F. There were 20 English word pairs in each of three conditions: transparent, opaque and form word pairs. The transparent list featured prime-target pairs that shared the same English base word and also were semantically related in meaning (e.g. sailor-SAIL) and the prime had an English suffix—i.e., the prime was a derivation of the target. The opaque list featured word pairs that shared the same English base orthography but were not related in meaning (e.g. corner-CORN) and the prime had an English suffix —i.e., the prime was pseudo-derived. The form list consisted of a prime word that also shared some of the same orthography of the target word (e.g. brothel-BROTH) and contained no English suffix —i.e., the prime word was not derived from the target. Each word target also appeared in another pair with an unrelated word prime that did not share any semantic or orthographic features with the target. Two lists were created to make sure no target word was shown twice during the task. All related and unrelated word primes were matched for frequency and length (Li et al., 2017).

During the task, participants also saw 60 English prime-nonword target pairs to create a lexical decision during the task. These distractor pairs were not provided in the Li et al. (2017) supplemental materials so they were created for this experiment. To create the nonwords, English primes were chosen from the MRC Psycholinguistic Database (Coltheart, 1981) by finding words of similar length and with similar suffix and monomorphemic endings to the experimental primes. Nonword targets were created by taking the base word of the MRC English prime word and changing one letter in the base word (e.g. limber – LEMB). The strategy to create the nonwords was adapted from Li et al. (2017). Two-thirds of the English distractor primes featured an English suffix (e.g. -er) and one-third featured an English monomorphemic ending (e.g. -el) to match the experimental pairs. Appendix F shows the complete word list for both List A and List B.

Because the experimental targets were paired with both related and unrelated word primes, two words lists, List A and List B, were created so that one list received the target with the related prime and one received the same target with the unrelated prime. Which list received the related or unrelated prime the target was alternated within each condition. Participants were assigned to either list A or list B upon arrival and the presentation of all prime-target items, experimental and distractor pairs, was randomized for each participant. All participants saw the same ten practice prime-target pairs before beginning the experimental portion or the task. Examples of prime-target word pairs assigned to List A and List B are included in Table 1.

For the masked priming task, each trial consisted of a forward mask of six hashmarks (e.g. ######) that was presented for 500 milliseconds (ms) followed by the prime in lowercase letters for 50 milliseconds, followed by the target word in capital letters for 1500 milliseconds. In order to match the computer monitor's refresh rate and show the prime for exactly 50 milliseconds, prime trials were programmed for 37 milliseconds.

Table 1. Examples of prime-target pair assignments to each list based on relatedness. R=related prime and UR=unrelated prime.

WORD TYPE	LIST A	LIST B
English word (UR/R)	trainer-PROPER	properly-PROPER
English word (R/UR)	sailor- SAIL	editor-SAIL
Nonword (R/UR)	factor-FECT	author-FECT
Nonword (UR/R)	poorly-SONSE	sensor-SONSE

Participants saw a white screen with black text in the center of the screen. Words were shown in Courier font, text size 24. Since the prime varied in word length, the forward mask did not always cover the prime completely. The forward mask was shorter than the prime word for 20 of the related prime words and 32 or the unrelated prime words.

Li et al. (2017) did not give details on whether the forward mask length varied based on the prime length or if it was the same length for all prime words. I chose to keep the prime length the same for all target words in order to be as consistent as possible across trials. Participants were asked if they could read any prime word after the experiment and any participants who said they could read a prime word were excluded from the analysis.

Participants were instructed to look at the screen and respond yes or no to whether the word on the screen was an English word. Participants made their selection on the keyboard using their dominant hand by pressing either the “G” or “H” key. A *yes* label was placed over the “G” key and a *no* label was placed over the “H” key. Participants were not given information about the mask or the prime but were asked three questions after completing the task to check if anyone was able to see the word prime. The first question was if participants saw anything after the hashmarks but before the target word, the second was if the participant could read what they saw (if anything) and the third question was if they saw any patterns in what they saw. Positive answers to these questions were used to eliminate participants from later analyses. Li et al. (2017) also asked these questions in their article but reported that no participants were able to see the word prime.

Data Analysis

Participants were eliminated from the various tasks as follows: Data from four participants were excluded from all tasks because they reported themselves as non-native English speakers on the language history questionnaire. All participants who completed the reading span task earned at least 80% accuracy on the grammaticality judgments, so no participant was eliminated from analyses for the reading span task. Lexical decision data was thrown out for the first nineteen participants because the instructions for the task were changed midway through the study. Masked priming data was thrown out for one participant because they pressed incorrect keys on the keyboard. Eight participants stated that they were able to read the word primes in the priming task and so their data was also excluded from the analysis on the priming task. These exclusions left 55 participants to be included for the lexical decision task analysis and 61 participants to be included for the masked priming task analysis.

In addition, four vocabulary words on the English vocabulary test were excluded because their average accuracy score was less than 10% correct (fecund, *M*=.03; redolent, *M*=.05; sobriquet, *M*=.05; bellicose, *M*=.07). Seven words were excluded from the English word analysis from the lexical decision task for two reasons. First, three of the low-performing words were words frequent in British vocabulary but not American English and they included the words W+S+ cookery, *M*= .51, W-S+ hoover, *M*= .73 and W+S- barrow, *M*= .73. The second set of excluded words included four words that earned less than 80% accuracy. The low performing excluded English words are W+S+ bony, *M*=.53, W-S- docile, *M*=.56, W-S+ mirage, *M*=.60, W+S+ duster, *M*= .78. I chose to remove these items because it was important to analyze words I was confident the sample population knew. I wanted the first analysis to serve as a reliable baseline of word response behavior for the non-native and elementary age groups in the later experiments. An anova test showed word length and word frequency averages between lists did not significantly differ after these words were removed (length, $F = 1$, frequency, $F < 1$).

Three nonwords from the lexical decision task were excluded for earning less than 50% average accuracy (W+S+ peacher, *M*=.39; W+S+ clockage, *M*=.39; W+S+ proudy, *M*=.46). All of these words were in the W+S+ condition which is the most English-like condition of the four conditions. Also, two of these words sound very similar to English words (peacher/preacher and proudy/proudly) and it is likely this is why they were selected as English words by the majority of the sample.

The statistical analyses applied to the lexical decision task and the masked priming task included a repeated measures ANOVA and mixed effect logistic regression for accuracy and a repeated measures ANOVA and linear mixed effect model for response times. All repeated measures ANOVA tests used both subject and item analyses. ANOVA tests were computed using SPSS software (IBM Corp, 2017).

Mixed effect models were computed in R (R Core Team, 2017) using the package lme4 (Bates, Machler, Bolker & Walker, 2014). A maximal model approach was selected for all linear mixed effects models for this experiment. This approach was chosen based on previous studies that have applied linear mixed models for lexical decision tasks (Li et al., 2017) and also to avoid an increase in Type 1 error present in random intercepts only models (Barr, Levy, Scheepers & Tily, 2013). In addition, using a maximal model approach for linear mixed effects models is highly recommended in order to fully account for variability present in the random effect variables for most psycholinguistic experiments (Clark, 1973). Random slope parameters were selected based on the hypotheses and design of the particular task for both the lexical decision and masked priming task. Models with and without interaction terms were compared using a log-likelihood goodness of fit test. The model that explained the data significantly better than the other was the one used for the results. This was almost always the model with interaction terms although not always. If a model with interaction terms was not the best model to explain the data, the model without interactions was used. The exact model used for reporting results is indicated before results are given.

One consequence to the maximal model approach is that it does significantly decrease statistical power (Matuschek, Kliegl, Vasishth, Baayen and Bates, 2017). In fact, maximal model parameters in linear mixed effects testing are notorious for producing issues related to quality of model fitness and one recommendation for a better model fit is to only enter random intercepts for random effects (Brauer and Curtin, 2017). Any issues related to model fit and convergence are listed in the results section. For the results in this dissertation, model fit improved for all models after response times were inverse log transformed to correct for positive skewness in the distributions. A description of the full statistical models and any excluded participants are included before the results are reported for each task. The statistical output from all linear mixed effect models are included in Appendix G. The two statistical models for accuracy and response times will be compared in the results and discussion sections to evaluate which better explained the data.

The vocabulary and reading span test were also correlated with participant performance on the experimental tasks using a Pearson Correlation test.

Results

English Vocabulary Test

Seventy-four participants completed the vocabulary test during our experiment and, after I removed the four low accuracy items, the average score was 54% accuracy with a standard deviation of 11%. Each vocabulary item had five answer options and chance was 20% accuracy. A one-sample t-test shows that the group performed significantly above chance, *t* (73) = 40.99, *p*

< .001, 95% CI, .51 - .56. The average response time to choose a vocabulary definition was 6.39 seconds with a standard deviation of 2.17 seconds.

For the ANOVA models, vocabulary performance scores were divided into two groups based on their z-scores values. Z-scores below 0 were assigned to the low vocabulary proficiency group and Z-scores above 0 were assigned to the high proficiency group. For the mixed effects models, vocabulary group was entered as a fixed factor in the models.

Reading Span Test

Forty-three participants completed the Reading Span test of the seventy-four participants total. The average partial score was 16.19 with a standard deviation of 8.10 points.

Lexical Decision Task

Data from 55 participants were included in the statistical analyses. Tests were conducted separately for English words and nonwords as done in Casalis et al. (2015a). Accuracy was analyzed by applying a 2 (Base present/absent) X 2 (Suffix present/absent) X 2 (Vocabulary High/Low) mixed Factorial ANOVA test using SPSS (IBM Corp, 2017). For the subject analysis, the first two variables are within factors and vocabulary group is a between factor. For the item analysis, all three variables are between factors. Subject (F1) and Word Item (F2) were analyzed as random variables and minF′ was then computed (Clark, 1973). Accuracy was also analyzed using a mixed effect logistic regression model in the lme4 package (Bates et al., 2014) in R (R Core Team, 2017).

For all lexical decision task mixed effects analyses in this experiment, the base, suffix and vocabulary group variables were entered as fixed effects. Subject and word item variables were entered as random intercepts and random slopes were always selected for the subject random effect since the word random effect never had a repeated factor in the design. Random slope parameters are specified before the results are discussed in each section. Significant effects for accuracy are reported in z-values and significant effects for response times are reported in t-values to match the reporting style exhibited in Li et al. (2017). P-values are given next to significant z and t values. For logistic models, p-values were taken from the summary() function which uses the Wald z-test. The lme4 package does not provide p-values in the summary() function for linear models so p-values were computed using the Anova() function from the car package (John Fox and Sanford Weisberg, 2011) which uses the Walds chi-square test.

Incorrect responses were excluded from the response time analysis and any response slower than 2.5 standard deviations above the mean per subject were also excluded. This removed 1.2% of the English word responses and 3.2% of the Nonword responses. Table 2 shows the average error rate and response times for the English and nonword categories separated by vocabulary group.

Response times were first analyzed by applying a 2 (Base present/absent) X 2 (Suffix present/absent) X 2 (Vocabulary High/Low) ANOVA test using SPSS (IBM Corp, 2017). As stated above, the first two variables are within factors and vocabulary group is a between factor for the subject analysis and all three variables are between factors for the items analysis. Subject (F1) and Word Item (F2) were analyzed as random variables and minF′ was then computed (Clark, 1973). A linear mixed effects model was also tested using the same parameters as the accuracy model.

English Word Accuracy

There was no main effect for Base, $F_1 < 1$, $F_2 < 1$ or Suffix, $F_1 < 1$, $F_2 < 1$. There was a main effect of Vocabulary group for items but not subjects, F_1 (1, 53) = 1.38, $p = .24$, F_2 (1, 138) = 5.29, $p < .05$, $\eta_p^2 = .04$.

Table 2. Adult Native Speaker means and standard deviation values for accuracy and response times on the lexical decision task.

	W+S+	W+S-	W-S+	W-S-
English Accuracy (%)				
Low Vocabulary	90.18 (5.85)	98.21 (2.44)	93.04 (5.15)	95.88 (4.10)
High Vocabulary	93.33 (5.72)	96.48 (3.44)	93.89 (5.06)	98.15 (3.15)
English RT (ms)				
Low Vocabulary	714.53 (81.18)	680.59 (82.39)	709.11 (84.06)	688.16 (82.28)
High Vocabulary	699.63 (79.12)	674.35 (82.22)	697.46 (73.47)	683.22 (74.61)
Nonword Accuracy (%)				
Low Vocabulary	73.04 (16.12)	88.75 (9.49)	91.96 (9.56)	90.54 (7.86)
High Vocabulary	76.48 (1.41)	94.07 (7.60)	95.19 (5.63)	94.07 (5.72)
Nonword RT (ms)				
Low Vocabulary	926.62 (134.83)	856.58 (151.93)	812.43 (118.55)	803.07 (131.04)
High Vocabulary	847.82 (91.82)	802.80 (102.68)	788.66 (114.32)	792.67 (101.16)

The interaction between the Base and Suffix variable was significant, minF′ (1, 181) = 5.73, $p < .05$, (F_1 (1, 53) = 15.03, $p < .001$, $\eta_p^2 = .22$, F_2 (1, 138) = 9.27, $p < .01$, $\eta_p^2 = .06$). When there was no base word, adding a suffix to the target word improved accuracy ($M_{subjects}$=.99, *SE*=.004, M_{items}=.99, *SE*=.007) compared to when there was no suffix ($M_{subjects}$=.97, *SE*=.005, M_{items}=.97, *SE*=.006), minF′ (1, 124) = 4.00, $p < .05$, (F_1 (1, 54) = 9.80, $p < .01$, $\eta_p^2 = .15$, F_2 (1, 70) = 6.70, $p < .05$, $\eta_p^2 = .09$). When there was a base word in the target word, adding a suffix decreased accuracy ($M_{subjects}$=.97, *SE*=.006, M_{items}=.97, *SE*=.007) compared to when there was no suffix ($M_{subjects}$=.99, *SE*=.003, M_{items}=.99, *SE*=.006) but this was only significant by subjects not items (F_1 (1, 54) = 5.02, $p < .01$, $\eta_p^2 = .09$, F_2 (1, 72) = 2.83, $p = .097$).

This pattern is the opposite of what I expected as I assumed English words with replicable morphemes (suffixes like -er and -ly) would be more familiar to the subjects and therefore receive higher accuracy averages. In this case, when a base word is in the word adding a suffix to the word removes the accuracy advantage present in the no base word condition. This interaction is depicted in Figure 1.

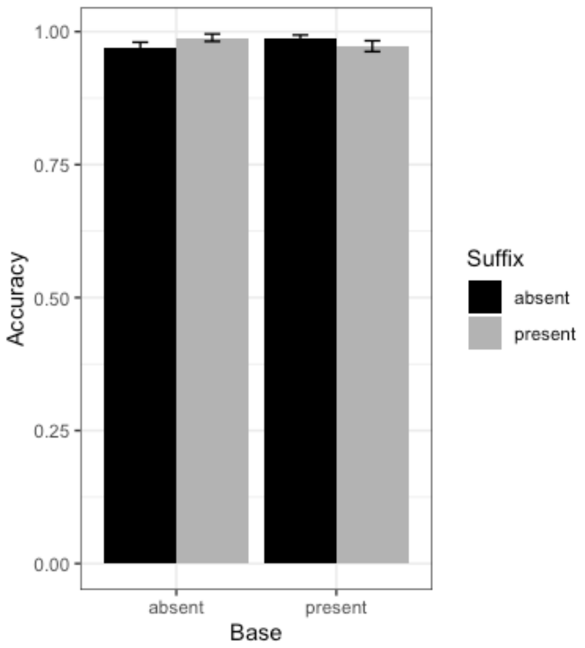


Figure 1. Average accuracy for base and suffix variables on English target words in native speaking adults.

The three-way interaction between Base, Suffix and Vocabulary Group was significant for both subjects and items but the minF′ was not significant, minF′ (1, 158) = 2.84, p =.094 (F_1 (1, 53) = 5.88, $p < .05$, $\eta_p^2 = .10$, F_2 (1, 138) = 5.48, $p < .05$, $\eta_p^2 = .04$).

A mixed effect logistic regression was conducted using the binomial link function to test accuracy. A base, suffix and vocabulary group interaction was entered as a fixed effect because my hypothesis states that these factors will interact. Subjects and word items were entered as random effects and were given random intercepts. A base by suffix interaction was assigned as the random slope for the subject random effect but this did not produce a fitted model and I received a convergence warning in the output. The observations in the dataset did not produce enough variance under the model parameters to accurately rely on the data in the summary output. In other words, the model parameters were too complex for the dataset.

A simpler model was selected based on the design of the experiment and the preliminary results from the ANOVA. The subject random effect was given a random slope for suffix. This also did not converge. I also entered only random intercepts for the random effects terms and this also failed the convergence test. I then separated the random effect terms into two models so that the first model tested subject random effects and the second model tested item random effects. This is the most complex model that would converge (or fit).

The reduction in model terms was guided by the advice in Brauer and Curtin (2018) for when mixed effect models to do not converge. They state that if a model does not converge, to simplify the model parameters until it reaches a better fit. Whenever, I received a convergence warning after entering a model, I used a convergence test to recheck the model fit and made decisions based on whether the values from the convergence test were above or below the < .001 threshold. The script for this convergence test is included in Appendix G and was sourced from an online thread on R addressing convergence issues related to LME model fit. The link to this thread is also in Appendix G. All models used in the results reported in this paper were below the recommended convergence threshold.

Before I simplified any models in the results, I tried several methods recommended by Brauer and Curtin (2018). I centered the predictor variables, and this always produced a poorer model fit. I adjusted the response time distribution to an inverse logarithm distribution for all datasets which improved model fit but did not always guarantee a model to pass the convergence test. I removed the random intercepts terms which produced a poorer model fit. Therefore, when simplifying models in this results section and later results section, I chose to remove random slope terms that were not essential to the design of the experiment. Since my hypotheses focus on the behavior of the participant sample, I chose to prioritize random slopes for the subject random effect and not the word random effect when models did not converge.

Some models still did not converge after reducing the complexity of random slope terms and so random slopes were removed entirely. This increased the risk of Type 1 error (Brauer & Curtin, 2018) but produced converging models. I also had to conduct separate tests for each random effect on some simple effect tests because even the random intercept only model still did not converge. Any time a mixed model was conducted and reported in this and later results sections, I report the exact model that was entered and explain why it was entered in that manner.

The model results for English word accuracy did not match the ANOVA results in that there were main effects for base, $z_{subjects}$=2.94, p <.01, z_{items}=2.61, p <.01 and suffix, $z_{subjects}$=2.78, p <.01, z_{items}=2.49 p <.05 and vocabulary group, $z_{subjects}$=3.23, p <.01, z_{items}=3.27, p <.01. When a base word was included in an English word, accuracy improved by 1.47 points on average. When a suffix was included in an English word, accuracy improved by 1.52 points on average. The high vocabulary group scored 1.21 more points than the low vocabulary group on average.

In addition, there was a significant interaction between base and suffix, $z_{subjects}$=-3.67, p <.01, z_{items}=-3.12, p <.01. All of the simple effects models for this dependent variable were fitted using random intercepts only as this was the most complex model that would converge. Like the ANOVA results, adding a suffix to a no base English word improved accuracy, which was significant, whereas adding a suffix to an English word with a base decreased accuracy but like the ANOVA results this decrease was not significant.

There was also a three-way interaction between base, suffix and vocabulary group, $z_{subjects}$=2.09, p <.05, z_{items}=2.12 p <.05. The following simple effects models also fit with random intercepts only but not random slopes. Only the low vocabulary group had a significant interaction between base and suffix, z=-3.00, p <.01. In the low vocabulary group, when there was no base word adding a suffix decreased accuracy by 1.52 points on average, z=2.60, p <.01.

When there was a base word, adding a suffix increased word accuracy but this difference was not significant, z=-1.50. This interaction is depicted in Figure 2.

Information worth mentioning is that when I entered the variables as fixed factors but not as an interaction, neither base nor suffix predicted word accuracy. This means that when an interaction is not included in the model, accuracy changes that occurred by adding a base or suffix to the word are less meaningful. When an interaction is included in the model, accuracy differences from these predictors are more meaningful. In an ANOVA test, factor means are compared to the overall model intercept to test significance. In a logistic mixed effects model, the comparison is like a regression where the factor coefficients are calculated in reference to the other factors in the model. This explains why there are very low F values in the ANOVA for the base and suffix effects but significant z-values in the mixed effects model.

English Word Response Times

There was a main effect of Suffix for subjects but not items, (F_1 (1, 53) = 4.76, $p < .05$, $\eta_p^2 = .08$, F_2 (1, 138) = 1.63, $p = .20$). There was no main effect for Base for subjects or items, $F_1 < 1$, $F_2 < 1$ and there was no main effect of vocabulary group, $F_1 < 1$, $F_2 < 1$.

There was a significant interaction between the Base and Suffix variables for subjects but not items (F_1 (1, 53) = 9.64, $p < .01$, $\eta_p^2 = .15$, F_2 (1, 138) = 3.51, $p = .063$). There was no interaction between Base and vocabulary group, $F_1 < 1$, $F_2 < 1$ or Suffix and vocabulary group, $F_1 < 1$, $F_2 < 1$. There was no interaction between Base, Suffix and vocabulary group, $F_1 < 1$, $F_2 < 1$.

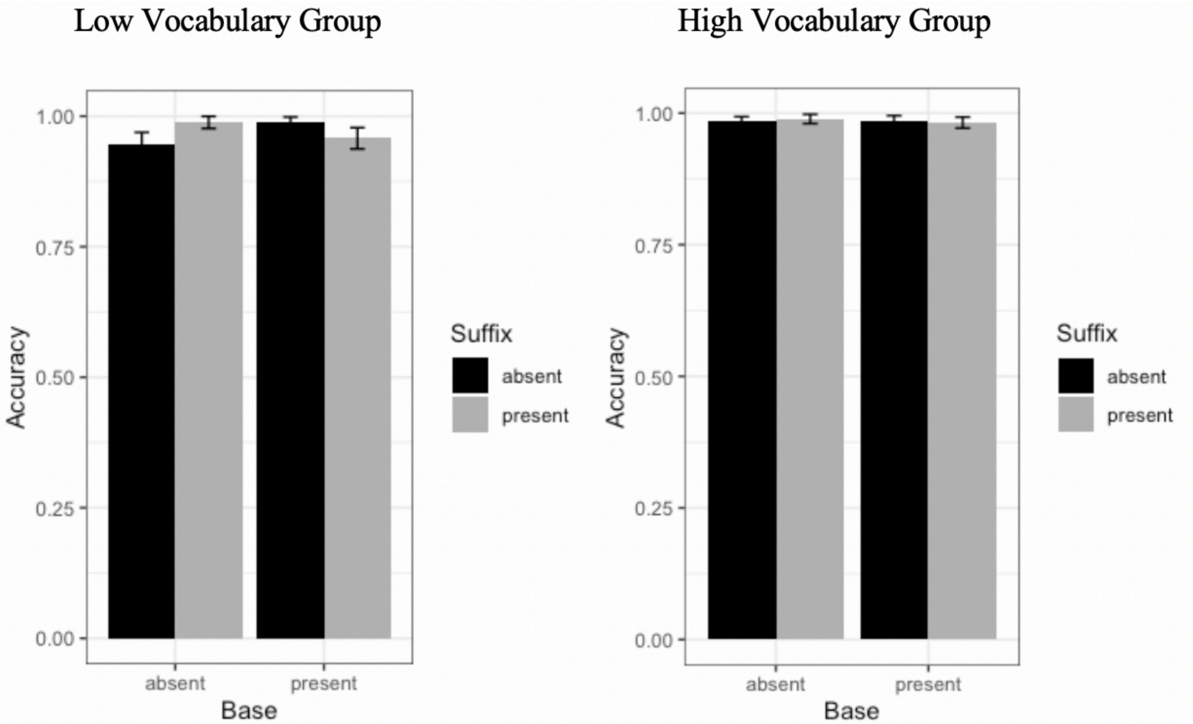


Figure 2. Average accuracy for base and suffix variables on English target words by vocabulary group.

Response times were also analyzed in a linear mixed effects model. An interaction between base, suffix and vocabulary group was entered as a fixed effect while the subject and word item variables were entered as random effects. The subject random effect was given a random intercept and random slope of base by suffix interaction and the word random effect was left as a random intercept. The random slope was chosen based on the ANOVA model which showed that subjects were significantly slower to respond to W+S+ condition words which suggests their behavior is driven by an interaction between the base and suffix variables.

The alternative model showed no main effect of suffix, *t* = -.02, base, *t* = .81 or vocabulary group, *t* = .18. There were no significant interactions between base and suffix variables, *t* = -1.16, base and vocabulary group *t* = .42 and suffix and vocabulary group, *t* = .31 and base, suffix and vocabulary group, *t* = -.03.

The results of the linear mixed effect model match the ANOVA in that none of the main effects or interactions are significant. The participants are not responding to target words differently based on the manipulations in the target word. They are also not responding differently based on their vocabulary proficiency group membership.

Nonword Accuracy

There was a main effect of Base, minF′ (1, 186) = 15.79, $p < .001$ (F_1 (1, 53) = 106.63, $p < .001$, $\eta_p^2 = .67$, F_2 (1, 146) = 18.53, $p < .001$, $\eta_p^2 = .11$). When participants saw a nonword without a base word, they were more accurate ($M_{subjects}$=.93, *SE*=.010, M_{items}=.93, *SE*=.011) than if they saw a nonword with a base word ($M_{subjects}$=.83, *SE*=.014, M_{items}=.86, *SE*=.012).

There was a main effect of Suffix, minF′ (1, 182) = 6.74, $p = .01$ (F_1 (1, 53) = 51.75, $p < .001$, $\eta_p^2 = .49$, F_2 (1, 146) = 7.75, $p < .01$, $\eta_p^2 = .05$). When participants saw a nonword without an English suffix, they were more accurate ($M_{subjects}$=.92, *SE*=.010, M_{items}=.91, *SE*=.011) than if they saw a nonword with an English suffix ($M_{subjects}$=.84, *SE*=.014, M_{items}=.87, *SE*=.012).

The main effects were qualified by a significant interaction between the Base and Suffix variable, minF′ (1, 179) = 11.27, $p = .001$ (F_1 (1, 53) = 93.60, $p < .001$, $\eta_p^2 = .64$, F_2 (1, 146) = 12.81, $p < .001$, $\eta_p^2 = .08$). When there was no base word, adding an English suffix to the nonword ($M_{subjects}$=.94, *SE*=.08, M_{items}=.93, *SE*=.016) made no difference in accuracy compared to the English suffix absent condition ($M_{subjects}$=.92, *SE*=.07, M_{items}=.92, *SE*=.016), F_1 (1, 54) = 1.52, $p = .22$, $F_2 < 1$. However, when there was a base word in the nonword, participants were significantly less accurate responding to nonwords with an English suffix ($M_{subjects}$=.75, *SE*=.151, M_{items}=.80, *SE*=.017) compared to the English suffix absent condition ($M_{subjects}$=.91, *SE*=.09, M_{items}=.91, *SE*=.016), minF′ (1, 95) = 13.88, $p < .001$ (F_1 (1, 54) = 96.11, $p < .001$, $\eta_p^2 = .64$, F_2 (1, 72) = 16.22, $p < .001$, $\eta_p^2 = .18$). This interaction is depicted in Figure 3.

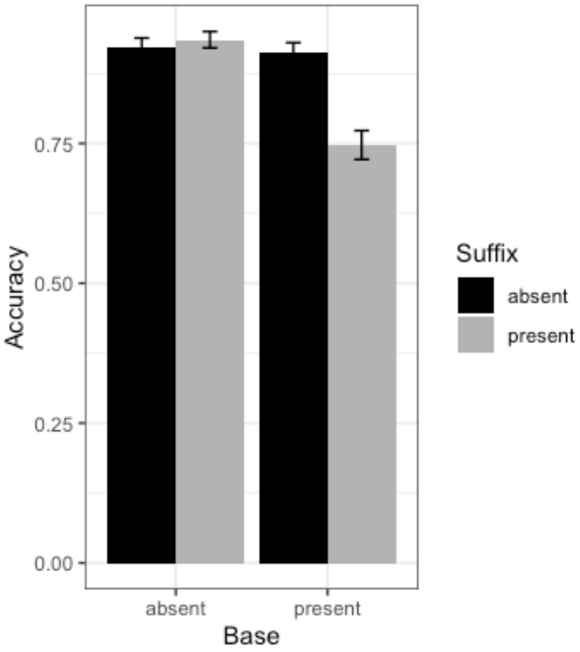


Figure 3. Average accuracy for base and suffix variables on nonword targets in native speaking adults.

There was no main effect of vocabulary group, F_1 (1, 53) = 3.26, p = .077, F_2 (1, 146) = 1.68, p = .20. There was no interaction between Base and vocabulary group, $F_1 < 1$, $F_2 < 1$ or Suffix and vocabulary group, $F_1 < 1$, $F_2 < 1$. There was no interaction between Base, Suffix and vocabulary group, $F_1 < 1$, $F_2 < 1$.

For the logistic mixed model, a base by suffix interaction was entered as a fixed effect and vocabulary group was entered as a fixed factor. Vocabulary group was not considered in the interaction because I wanted to simplify the model in order for it to converge. Previous tested models and the ANOVA showed that vocabulary group was not a significant factor for this dependent variable and so I chose not to include it in the interaction term. Subjects and word items were entered as random effects and were given random intercepts. A base and suffix interaction was entered as a random slope for the subject random effect but this model did not pass the convergence test. The suffix variable was selected as a random slope for the subject random effect as the base variable random slope failed to converge.

The model did not match the ANOVA results in that both the base, z=-.64 and suffix variables, z=.32 were not significant predictors of test accuracy. Vocabulary group was not a significant predictor of performance as also stated in the ANOVA, z=1.62.

There was a significant interaction between the base and suffix variable as a predictor, z=-3.11, p <.01. The simple mixed effect logistic regression models matched the simple effects ANOVA in that when a suffix was added to a word with a base word, accuracy decreased but when a suffix was added to a word without a base word, accuracy was equivalent to nonwords with no base word and no suffix. There were no other interactions.

The reason why the main effects were not present in the mixed effects model is because the model was considering the interaction term while calculating the factor coefficients. The ANOVA was evaluating the factors an interaction independently. The significant interaction but no significant main effects for the mixed effects model means that the participant behavior is explained by the interaction between the two variables and not by base or suffix alone. Looking at Figure 3 on the previous page, it is clear that the accuracy difference is in one condition and there are no main effects.

Nonword Response Times

There was a main effect of Base, minF′ (1, 197) = 25.10, $p < .001$ (F_1 (1, 53) = 79.90, $p < .001$, $\eta_p^2 = .60$, F_2 (1, 146) = 36.59, $p < .001$, $\eta_p^2 = .20$). When participants saw a nonword without a base word, they were faster to respond it was not an English word ($M_{subjects}$=799.21, SE=15.28, M_{items}=799.72, SE=7.16) than if they saw a nonword with a base word ($M_{subjects}$=858.46, SE=16.09, M_{items}=862.34, SE=7.47).

There was a main effect of Suffix, minF′ (1, 197) = 7.76, $p < .01$ (F_1 (1, 53) = 34.25, $p < .001$, $\eta_p^2 = .39$, F_2 (1, 146) = 10.03, $p < .01$, $\eta_p^2 = .06$). When participants saw a nonword without an English suffix, they were faster to respond that it was not an English word ($M_{subjects}$=813.78, SE=16.05, M_{items}=814.64, SE=7.16) than if they saw a nonword with an English suffix ($M_{subjects}$=843.88, SE=15.04, M_{items}=847.42, SE=7.47).

These effects were qualified by a significant interaction between the Base and Suffix variable, minF′ (1, 199) = 5.38, $p < .05$ (F_1 (1, 53) = 19.92, $p < .001$, $\eta_p^2 = .27$, F_2 (1, 146) = 7.37, $p < .01$, $\eta_p^2 = .05$). When there was no base word, adding an English suffix to the nonword ($M_{subjects}$=800.76, SE=116.36, M_{items}=802.06, SE=10.13) made no difference in the response time compared to the suffix absent condition ($M_{subjects}$=797.97, SE=116.36, M_{items}=797.38, SE=10.13), $F_1 < 1$, $F_2 < 1$. However, when there was a base word in the nonword, participants responded significantly slower to nonwords with an English suffix ($M_{subjects}$=887.94, SE=121.37, M_{items}=892.78, SE=10.99) compared to the suffix absent condition ($M_{subjects}$=830.18, SE=131.73, M_{items}=831.90, SE=10.13), minF′ (1, 180) = 11.79, $p = .001$ (F_1 (1, 54) = 47.72, $p < .001$, $\eta_p^2 = .47$, F_2 (1, 70) = 15.65, $p < .001$, $\eta_p^2 = .18$). This interaction is illustrated in Figure 4.

There was no main effect of the vocabulary group variable for subjects but there was for items, minF′ (1, 69) = 1.62, $p = .21$ (F_1 (1, 53) = 1.85, $p = .18$, F_2 (1, 146) = 13.07, $p < .001$, $\eta_p^2 = .08$). There was a significant interaction between the presence of a base word in the word item and the vocabulary proficiency of the participant for both subjects and items but the minF′ was not significant, minF′ (1, 198) = 3.21, $p = .075$ (F_1 (1, 53) = 13.78, $p < .001$, $\eta_p^2 = .21$, F_2 (1, 146) = 4.18, $p < .05$).

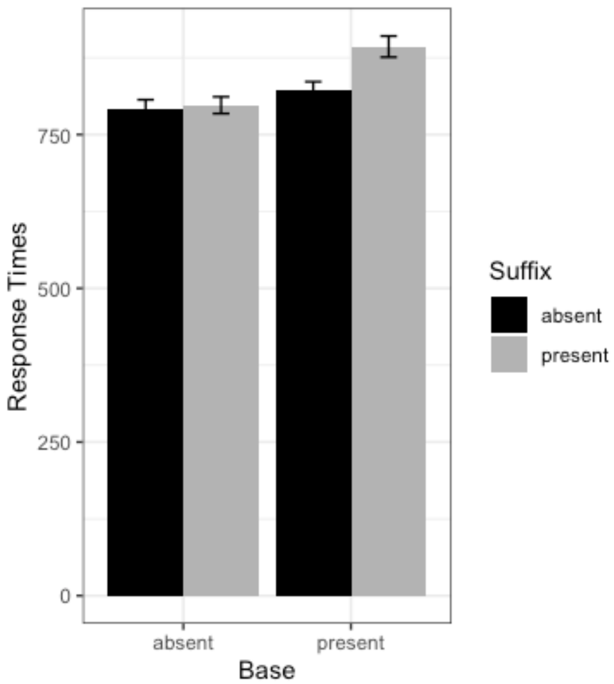


Figure 4. Average response times for base and suffix variables on nonword targets in native speaking adults.

Response times were also analyzed in a linear mixed effects model. A base, suffix and vocabulary group interaction was entered as a fixed effect while the subject and word item variables were entered as random effects. The subject random effect was given a random intercept and a base by suffix interaction random slope and the word random effect was left as a random intercept. The random slope was chosen based on the ANOVA results which showed that subject responses vary based on a specific base and suffix level interaction present in the target word. This model converged appropriately.

There was a main effect of base, $t = -2.14$, $p < .05$ in that when there was a base word present in the nonword, participants responded slower to the target word. There was no effect of suffix, $t = -.21$ or vocabulary group, $t = .44$.

There was significant interaction between base and suffix variables, $t = -2.86$, $p < .01$ and the simple effects pattern matched the ANOVA test in that there was no main effect of suffix when no base word was present in the nonword but participants were significantly slower to respond to nonwords that had a base word and a suffix compared to nonword that had a base word and no suffix. There was no significant interaction between suffix and vocabulary group, *t* = -.05, base and vocabulary group, *t* = 1.27 or base, suffix and vocabulary group, *t* = -.9.

The results from the linear mixed effects model and the ANOVA are the same except for there was no suffix main effect in the mixed model but there was in the ANOVA. This is most likely explained by the fact that the interaction explains the variability in the suffix variable in the mixed model whereas the ANOVA calculations are considering the means of the suffix variable against the model intercept and the interaction term means against the model intercept. The differences between the main effects in the ANOVA compared to the mixed models is showing that the mixed effects model is giving more specific guidance towards where the variability derives whereas the ANOVA shows the big picture in terms of what factors predict behavior and its up to the researcher to investigate and interpret using simple comparisons.

Masked Priming Task

Data from 61 participants were analyzed for English word trials only. Accuracy was analyzed with a 2 (Relatedness) X 3 (Condition) X 2 (Vocabulary Group) mixed Factorial ANOVA. The first two variables are within factors and vocabulary group is a between factor for the subject analysis whereas relatedness is a within factor and condition and vocabulary group are between factors for the item analysis. Subject (F1) and Item (F2) variables were computed as random effects variables and minF′ was then computed.

Incorrect responses were excluded from the response time analysis. Response times faster than 150 milliseconds and slower than 2.5 standard deviations above the mean were also excluded. This removed 1.7% of the responses. Table 3 shows the average response times and error rate for each prime and relatedness category.

Response times were first analyzed by applying a 2 (Relatedness) X 3 (Condition) X 2 (Vocabulary Group) mixed Factorial ANOVA. The first two variables are within factors and vocabulary group is a between factor for the subject analysis whereas relatedness is a within factor and condition and vocabulary group are between factors for the item analysis. Subject (F1) and Item (F2) variables were computed as random effects variables and minF′ was then computed.

For the mixed effects analyses, the relatedness, condition and vocabulary group variables were entered as fixed effects. Subject and word item variables were entered as random effects and random slopes were chosen for both subject and word random effects based on model fit. These choices are outlined before each results section.

Accuracy

Applying the Greenhouse-Geisser correction because sphericity was violated, there was a main effect of Condition but the minF′ was not significant, minF′ (2, 144) = 1.92, $p = .15$ (F_1 (2, 118) = 4.01, $p < .05$, $\eta_p^2 = .064$, F_2 (2, 54) = 3.69, $p < .05$, $\eta_p^2 = .12$). There was no main effect of Relatedness, $F_1 < 1$, $F_2 < 1$ and no interaction between condition and relatedness, F_1 (2, 118) = 1.15, $p = .32$, $F_2 < 1$. There was no main effect of vocabulary group, $F_1 < 1$, $F_2 < 1$. There was no interaction between Condition and vocabulary group, $F_1 < 1$, $F_2 < 1$, Relatedness and vocabulary group, F_1 (1, 59) = .2.77, $p = .10$, F_2 (2, 54) = 2.15, $p = .15$ and no three-way interaction, $F_1 < 1$, F_2 (2, 54) = 1.66, $p = .20$.

A mixed effect logistic regression was conducted using the binomial link function. A relatedness by condition interaction and vocabulary group were entered as fixed effects and subjects and word items were entered as random effects and were given random intercepts. Condition was also entered as a random slope for the subject random effect. This is the most complex model that would converge. The model results matched the ANOVA results in that no main effects or interaction effects were significant, and all z-values were below 1.

Table 3. Adult Native Speaker means and standard deviation values for accuracy and response times on the masked priming task.

	Form (castle/cast)	Opaque (corner/corn)	Transparent (hunter/hunt)
Related Prime Accuracy (%)			
Low Vocabulary	97.43 (5.60)	96.57 (5.91)	98.00 (5.84)
High Vocabulary	97.69 (5.14)	98.08 (4.02)	97.69 (5.14)
Related Prime RT (ms)			
Low Vocabulary	1265.95 (88.77)	1239.05 (81.78)	1202.80 (89.80)
High Vocabulary	1243.81 (71.90)	1256.94 (60.56)	1210.11 (49.22)
Unrelated Prime Accuracy (%)			
Low Vocabulary	98.29 (3.82)	96.57 (5.91)	98.57 (3.55)
High Vocabulary	97.69 (5.14)	95.38 (6.47)	98.08 (4.02)
Unrelated Prime RT (ms)			
Low Vocabulary	1262.93 (84.95)	1266.99 (80.48)	1265.59 (83.29)
High Vocabulary	1257.05 (60.91)	1264.11 (71.94)	1246.66 (67.25)

Response Times

There was a main effect of Relatedness, minF′ (1, 113) = 18.35, $p < .001$ (F_1 (1, 59) = 34.22, $p < .001$, $\eta_p^2 = .37$, F_2 (2, 54) = 39.57, $p < .001$, $\eta_p^2 = .42$). Participants were slower to respond to targets with an unrelated prime ($M_{subjects}$=1260.56, *SE*=9.01, M_{items}=1252.26, *SE*=6.67) compared to those with related primes ($M_{subjects}$=1236.44, *SE*=8.94, M_{items}=1226.81, *SE*=4.56). The subject and item effects were significant for Condition but the minF′ was not significant, minF′ (2, 81) = 2.67, p =.075, (F_1 (2, 118) = 13.73, $p < .001$, $\eta_p^2 = .19$, F_2 (2, 54) = 3.32, $p < .05$, $\eta_p^2 = .11$).

There was a significant interaction between Relatedness and Condition, minF′ (2, 149) = 5.96, $p < .05$, (F_1 (2, 118) = 11.80, $p < .001$, $\eta_p^2 = .17$, F_2 (2, 54) = 12.04, $p < .001$, $\eta_p^2 = .17$). A simple comparison showed that response times for by relatedness did not significantly differ for the form condition, $F_1 < 1$, $F_2 < 1$, but positive priming effects were significant for the opaque condition, minF′ (2, 94) = 6.63, $p < .01$, (F_1 (1, 74) = 7.82, $p < .01$, $\eta_p^2 = .10$, F_2 (2, 57) = 8.68, $p = .001$, $\eta_p^2 = .23$) and transparent condition, minF′ (2, 94) = , $p < .01$ (F_1 (1, 72) = 77.23, $p < .001$, $\eta_p^2 = .52$, F_2 (2, 57) = 8.68, $p = .001$, $\eta_p^2 = .23$). This means that participants responded to related opaque and transparent target words significantly faster than unrelated opaque and transparent target words. The priming effects for the transparent ($M = 49.27$, $SD = 47.90$) and opaque condition ($M = 18.65$, $SD = 56.98$) by subjects are significantly different than zero in a one-sample t-test, $t_{\text{transparent}}$ (72) = 8.79, $p < .001$, t_{opaque} (72) = 2.80, $p < .01$. These effects are depicted in Figure 5.

There was no main effect of vocabulary group for subject or items, $F_1 < 1$, $F_2 < 1$. There was no interaction between Relatedness and vocabulary group, F_1 (1, 59) = 1.55, $p = .22$, $F_2 < 1$ and Condition and vocabulary group, F_1 (2, 118) = 1.82, $p = .17$, $F_2 < 1$. There was no interaction between Relatedness, Condition and vocabulary group, F_1 (2, 118) = 2.98, $p = .055$, F_2 (2, 54) = 2.20, $p = .12$.

Response times were also analyzed in a linear mixed effects model. A relatedness, condition and vocabulary group interaction was entered as a fixed effect while the subject and word item variables were entered as random effects. The subject random effect was given a random intercept and relatedness and condition interaction random slope. The word random effect was given a random slope by relatedness since relatedness was a within factor in the experimental design for both random effects. These random effects were chosen based on the hypothesis of the experiment and the design of the study. This model did converge appropriately, and the results reported below are from the above fitted model.

The alternative model showed no effect of relatedness, $t = -1.70$, no effect for condition, $t = -.48$ and no effect for vocabulary group $t = -.50$. There was a significant effect of relatedness by condition, $t = 3.82$, $p < .001$.

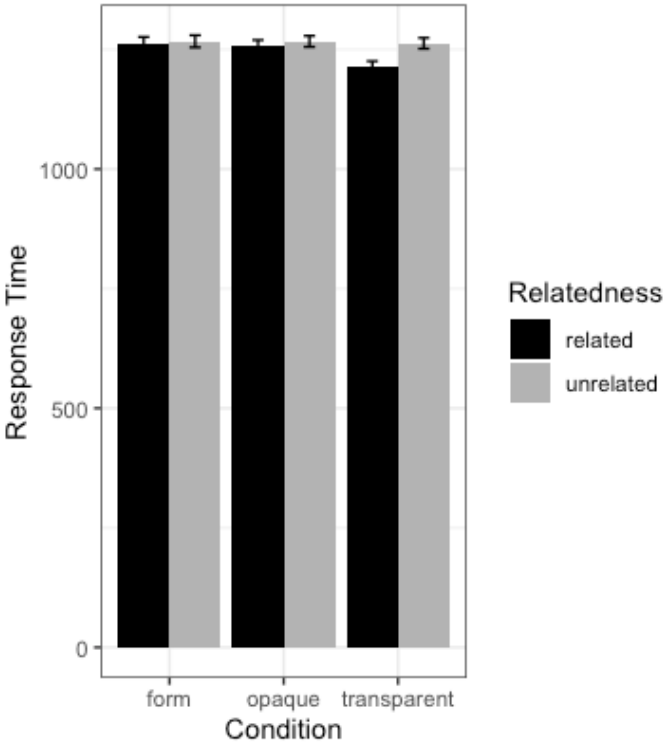


Figure 5. Average response time for related and unrelated conditions in native English-speaking adults.

Simple models were entered for each level of condition to investigate the whether any of the priming effects computed with the two relatedness types were significant. Both random effects were assigned relatedness as a random slope for these models. The results did not match the ANOVA tests in that significant positive priming effects were identified in the transparent, $t = 5.82$, $p < .001$ condition but the not the opaque, $t = 1.57$ and form condition, $t = .61$. In comparison, the ANOVA test found significant priming effects for both the transparent and opaque conditions.

There was no significant interaction of relatedness and vocabulary group, *t* = .42 and condition and vocabulary group, *t* = .85. There was no three-way interaction between relatedness, condition and vocabulary group, *t* = -1.05.

The results of the linear mixed effects model do not match the ANOVA test in a similar way to the previous comparisons. The interaction term seems to explain the variability in the relatedness and condition factors and so those factors are no longer significant in the model. Specifically, the area in which participant response time behavior is changing is in the related manipulation for the transparent condition as evidence by the simple comparisons. Regarding why the opaque priming effect is significant in the ANOVA test but not the mixed model, the ANOVA considers both the subject and item variability with the minF prime computation, but the mixed model also considers the slope of relatedness for both subjects and word items which is a component of the maximal model framework. This additional parameter could explain the difference between significance in the two tests. Barr et al. (2013) critiques models that do not include random slopes for within factors because of the increase in Type 1 error. The significant opaque priming effect on the ANOVA test could be an example of Type 1 error.

Vocabulary and Reading Span Correlations

A Pearsons bivariate correlation was conducted with vocabulary score and vocabulary response time against lexical decision and priming task performance. There were no significant correlations between vocabulary score and task performance but there were significant correlations between vocabulary response time and lexical decision response times. Those relationships are depicted below in Table 4. These correlations show a positive relationship between speed of vocabulary response and speed of lexical decision on the lexical decision task. Participants who were faster in their vocabulary answers, were also faster answering easier lexical decision items like the most-English like English words and the least-English like nonwords. This most likely means that participants with more confidence in their vocabulary answers were also more confident in their lexical decision answers for these conditions. Since accuracy is not a significant factor in the correlation for both the vocabulary test or the lexical decision test, I cannot say that vocabulary knowledge influenced word recognition responses.

There were no significant correlations between reading span scores and performance on the lexical decision or masked priming task. For this participant group, the cognitive skills used by participants during the shortened reading span task did not relate to the skills required for either the lexical decision or masked priming task.

Table 4. Pearson correlation between lexical decision response times and vocabulary test response times.

Lexical Decision Response Time	Vocabulary Response Time
W+S+ English Words	*r* = .314*, *p* = .02
W-S+ Nonwords	*r* = .330*, *p* = .014
W-S- Nonwords	*r* = .391**, *p* = .003

Discussion

Adult native speakers of English completed an English vocabulary test, the shortened version of the reading span test, a lexical decision task and a masked priming task. My hypotheses before the experiment were that, first, adult native speakers of English would be sensitive to word structure changes in the lexical decision task as reflected in their accuracy and response times. This would provide evidence the group is mentally decomposing complex words during word processing. My second hypothesis was that adult native speakers would engage in obligatory decomposition in the masked priming task by producing significant priming effects in the opaque prime-target pair condition (e.g. corner-CORN). My last hypothesis was that a participant's vocabulary test score would influence performance on the two experimental tasks.

Lexical Decision Task

The results supported my hypothesis for the lexical decision task. Participants did respond differently both in accuracy and response speed when the structure of the target word changed for both base and suffix morphology however these patterns were not the same for both English and nonwords and were not consistent across dependent measures.

For English target words, the mixed effects model showed that adding a suffix to a base absent word improved accuracy but adding a suffix to a base present word did not affect accuracy. This shows that having no English morpheme cue in the word decreases lexical decision accuracy but having two cues, both a base and suffix morpheme in the same does not provide an accuracy benefit. The W-S+, W+S- and W+S+ all perform the same in terms of accuracy and they all have at least one English morpheme in the word. The odd condition in the results is the W-S- condition which has no English morpheme in the words in that list.

The base and suffix interaction in the accuracy data is in some ways the opposite of what I would expect. Adding a suffix to a word should improve accuracy in all situations not just in the no base condition. It is possible that the benefit of having a base word in the word washed out any added benefit of adding a suffix to a base present word. In other words, the reader is looking for the base word to make a lexical decision and not the suffix, but they will use the suffix structure if it is there without a base word. Another explanation is that because nonwords were mixed in with English words in this task it is possible the participants viewed the words in the W+S+ condition as trick words or nonwords made to look like English words due to their extra English-like form. This would have decreased the accuracy in this condition and made it look like the participants were not receiving an accuracy benefit from having two English morphemes instead of one.

Another explanation for why this pattern occurred could be the amount of time participants had to respond to the target word. Participants were allowed 5 seconds to respond in the lexical decision task compared to 1.5 seconds in the masked priming task which was completed first in the experimental procedure. This extra time may have created more room for participants to doubt their initial instincts.

The accuracy model also showed a significant three-way interaction between base, suffix and vocabulary group which highlighted how vocabulary proficiency influenced target accuracy. Low vocabulary participants benefited from a suffix in the target word when there was no base word but did not benefit from a suffix being present when there was a base word. This means low vocabulary readers rely on an English morpheme cue more than the high vocabulary readers. Both the ANOVA and the linear mixed effects model revealed no significant effects in the response time data. This means that the response time behavior from the group did not vary based on the manipulations in the word or the vocabulary proficiency of the participant. This is most likely because the target words are easy enough to produce a whole word access strategy for word recognition instead of morphological decomposition. It is also possible that the words are so quickly accessed that any morphological decomposition variability is difficult to detect.

Given that this task is so easy for the participant group, it is difficult to identify whether a lack of results provides proof that morphological decomposition is not occurring or whether the words are just quickly being morphologically decomposed and so no response time differences are detected. Because there is no semantic manipulation between the conditions, this task does not identify obligatory decomposition, it only identifies whether morphological decomposition is occurring and whether that strategy is efficient or not efficient for lexical recognition. In this case, I can not confirm or disconfirm that morphological decomposition is occurring but whatever strategy the participant sample is using, whole word access or morphological decomposition, it is equally efficient across conditions.

Since the English word effects come from the accuracy data, I can only evaluate these patterns in terms of participant word knowledge. English words in the W-S+ were more easily recognizable as English words for participants overall and especially low vocabulary participants. It is unclear how the structure of the English words influenced word processing as indicated by response times. The lack of response time effects may be due to the data points removed from the seven exclusion words that earned below 80% accuracy. However, this exclusion was important to create a clean study design.

The results for the nonword targets are more readily interpretable than the English target word results and also gives information about morphological decomposition during visual word processing. The results show that mental decomposition occurs even when a reader has never encountered the letter string before. The accuracy and response time ANOVAs showed a main effect of base and suffix and a significant interaction between the two variables. When a nonword had no English morphemes (e.g. spetle) and when a nonword had no base word but an English suffix (e.g. hosper), performance was equivalent on average. However, when nonwords with a base word morpheme and no suffix (e.g. hillet) were compared to a nonwords with both a base word and suffix (e.g. ruly), there is significantly better accuracy and faster response times for the base word/no suffix nonword. This means that when processing unknown words, adding a suffix to a word without a base word doesn’t create a more English-like nonword. However, for words with a base word, adding a suffix suddenly makes the nonword more likely to be treated as an English word.

The participants in this experiment are using the English structure information to make judgments about the viability of the nonword as an English word. If participants were not using this information, accuracy and response times would be equivalent across conditions. It makes sense that participants are using a morphological decomposition strategy instead of a whole word retrieval strategy for unfamiliar words since they have never encountered the word before, they have no lexical item to retrieve from memory. In this way, the participants are relying on bottom up processing versus top-down processing (Andrews & Bond, 2009)

Unlike the English target analysis, vocabulary group membership did not influence accuracy or response times for nonwords. This means that participants of varying vocabulary proficiencies take advantage of the same strategies when trying to make lexical decisions for nonwords that include English morphemes. Additionally, reading span performance did not correlate with any dependent measure in this group which indicates morpheme segmentation does not require the same cognitive process as a reading span task.

The lexical decision task in a native English-speaking adult population shows that the base and suffix morpheme may not be equivalent in their role for English complex word processing. By showing an accuracy benefit for the W-S+ condition compared to the W-S- condition but no benefit for the W+S+ condition compared to the W+S- condition, there are two possibilities. First, the base morpheme is the preferred cue for English word recognition or only one morpheme cue is needed to make a decision and there is no added benefit to having both a base and suffix morpheme in the word. The interaction between vocabulary group and base word and vocabulary group and base and suffix shows that whichever is the best explanation, the strategy is particularly useful for low vocabulary readers.

For nonwords, having one English morpheme present is not significantly more confusing than having no English morphemes present but having both a base and suffix present in the nonword leads to significantly lower accuracy and slower response times than the other three conditions. This means that when trying to process a new or unfamiliar word, readers use both the base and suffix to make their decisions regarding what is and is not a nonword. This strategy is also inefficient, and costly in terms of response times, when making a correct rejection. This could be a sign of lack of confidence or extra search time using the English morphemes in the word to determine whether it is in fact a word or not.

Some research has suggested that readers process nonwords by mentally removing and examining the base form of the word and not the prefix or the suffix. This is known as base word stripping (Grainger & Beyersmann, 2017) as opposed to affix stripping (removing the prefix or suffix). However, this is not a strategy the participants are using during this task since there is no main effect of base word. Since the W+S+ nonword condition is the poorest performing condition, the subjects are examining both the base word and the suffix structure in the nonwords.

Masked Priming Task

The response time results from the ANOVA tests indicated that priming effects occurred in both the opaque and transparent conditions. When related primes from these conditions were subtracted by unrelated primes in these conditions, the related primes produced significantly faster target response speeds compared to the unrelated primes. This means that native English speakers in this sample mentally decomposed falsely derived word pairs like corner-corn. The obligatory decomposition theory (Rastle & Davis, 2008) is supported by these results because there are priming effects in both the transparent and opaque word pair conditions. The opaque priming effect was not significant in the mixed model most likely due to item variability with the opaque word list. More research is needed to understand whether a more controlled opaque word list for the participant population would also produce significant positive priming effects.

Another important question for later research is why the priming effect in the opaque condition is weaker compared to the transparent condition when participants are theoretically using the same decomposition strategy. Overall, the results in this experiment replicate the general priming pattern present in masked priming studies reviewed in Rastle and Davis (2008).

There were no significant effects for accuracy. This is most likely because the participants recognized the targets words consistently across conditions. For the ANOVA response time analysis, participants responded faster to related targets with related primes compared to unrelated primes and they also responded to targets with transparent primes faster than the other pair categories. The linear mixed effects model only found a significant interaction between relatedness and condition and this interaction was defined by the significant priming effect in the transparent condition.

Unlike the previous task, vocabulary proficiency was not a factor in how participants responded to the manipulations in this task. Participants with low vocabulary proficiency and participants with high vocabulary proficiency responded to the target words in the same manner. Some research has shown that individual differences in vocabulary proficiency do influence morphological decomposition strategies on masked priming tasks (Beyersmann, Casalis, Ziegler & Grainger, 2015; Schmidtke, Van Dyke & Kuperman, 2018) but the results from this experiment do not support that research. Because readers are using their knowledge of English orthography as well as vocabulary and morphology to complete this task, some research has suggested spelling ability is a better predictor of masked priming differences between subjects (Andrews & Hersch, 2010).

The two statistical models used for the analyses, the ANOVA tests and the mixed effects models, provided the same information in different ways with some exceptions. When there was a significant interaction, the mixed model reported that the main effects were not significant except for the interaction. This occurred when there were main effects identified in the ANOVA tests for the same dependent variable. This is most likely because the mixed model calculations are considering the factor means in relation to other factors entered in the model and the ANOVA is comparing each factor means to the intercept individually.

Exceptions to this include the English target accuracy results and the masked priming response times. The opaque priming effect being present in the ANOVA but not the mixed model is an example of the maximal model approach eliminating type 1 error. However, the English accuracy results are a bit trickier. The English target accuracy results showed main effects for base and suffix when the ANOVA showed F values less than 1 for each of those factors. I configured the parameters for that model multiple ways and re ran the best converging model multiple times and I still got the same results. This is an example of the importance of understanding how mixed effects models identify and report variability in comparison to ANVOA tests. My thoughts are that when all of the coefficients were considered together in the mixed model, important variability in the main factors were revealed that were not detected in the ANOVA test.

Overall, I find the mixed effects model approach better than the ANOVA tests mainly because it gives more specific information regarding where the variability is located in the data frame. I also feel like I have more control over how the model looks for variability given the hypotheses and design of my task. Finally, I believe the mixed effects models have more statistical power and are able to amplify effects that are hidden in the ANOVA test. This advantage was shown when the minF prime was not significant for a couple interactions in the experiment, but the mixed effects model interactions were. For these reasons, in the next two experiments, I will compute both minF prime and mixed effects model statistics for the dependent measures, but I will only report the ANOVA test results if it differs from the mixed model test.

Chapter 4. Experiment 2. Adult Non-Native Speakers

The second experiment replicates the experimental procedure from Experiment #1 but includes an adult non-native English-speaking participant group. My research questions are, first, whether non-native speakers will be sensitive to word and nonword base and suffix structure changes in the lexical decision task through their accuracy and response time behavior. My second question is whether non-native English-speaking adults show priming effects in both the opaque (corner-CORN) and form (castle-CAST) condition (overuse morphological decomposition) and whether this pattern is consistent across age of acquisition groups. Following previous literature and the studies from which these materials were adapted, I hypothesize non-native speakers will be sensitive to word structure in the lexical decision task and use both base words and suffix structures to improve the accuracy of their lexical decisions. For the masked priming task, I hypothesize this group will show priming effects in all three priming conditions: transparent, opaque and form and that participants who acquired English earlier in life will show less priming in the form condition.

Method

Participants

Participants were recruited from the Psychology Research Participant Pool and International Student English classes on the university campus. Five participants were recruited by word of mouth. None of the participants identified themselves as having a learning disability or language related disorder.

For this experiment, I was able to recruit 49 participants. Twenty-nine of the participants were females. All participants identified themselves as being proficient in two or more languages, one of which is English and in all cases English was acquired after birth. All participants reported having received at least a high-school education and all of the participants were currently enrolled in either an undergraduate or graduate program at the university. The average participant age was 22.54 years (*SD*=.4.46). The average age of English acquisition was 6.39 years of age (*SD*=4.41).

The language history questionnaire asked each participant to give their age of acquisition and proficiency in listening, speaking, reading and writing for each language they spoke including English. Their English age of acquisition and English proficiency ratings for each category are shown in Table 5. Proficiency was self-rated on a 1-7 scale (1=very poor, 2=poor, 3=limited, 4=average, 5=good, 6=very good and 7=excellent). Participants also gave their country of origin and their self-rated language learning ability using the same scale (very poor to excellent). English age of acquisition in each English mode and self-reported proficiency in each mode were all highly correlated with one another.

Represented countries of origin (and territory) included Bangladesh n=3, China n=6, Colombia n=2, Guinea n=1, Honduras n=1, India n=2, Japan n=3, Jordan n=1, South Korea n=1, Malaysia n=1, Mexico n=2, Nicaragua n=1, Pakistan n=1, Puerto Rico n=1, Russian n=1, Thailand n=1, Ukraine n=1, United Arab Emirates n=1, the United States of America n=12 and Vietnam n=4. Four participants did not report their country of origin. Participant average language learning ability was 4.62 (SD=1.09) which translate to a self-rating of “average”.

Represented spoken languages in the sample population included Spanish n=13, Chinese =8, French n=5, Vietnamese n=4, Bengali n=4, , Russian n=3, Hindi n=3, Arabic n=2, Italian n=2, Thai n=1, German n=1, Pashto n=1, Marathi n=1, Malay=1, Punjabi n=1, Japanese n=1, Sanskrit n=1, Latin n=1, Ukrainian n=1, Korean n=1, Gujarati n=1, Urdu =1.

Table 5. Average age of acquisition and self-reported proficiency ratings for each mode of English.

English Age of Acquisition and Self-Rated Proficiencies	Age (SD) and Mean Proficiency (SD): 1-7 Scale - Poor to Excellent
Age of Acquisition-Listening	5.91 (4.02)
Age of Acquisition-Speaking	7.86 (5.36)
Age of Acquisition-Reading	7.33 (3.60)
Age of Acquisition-Writing	7.50 (3.93)
Proficiency-Listening	5.82 (1.09)
Proficiency-Speaking	5.29 (1.19)
Proficiency-Reading	5.62 (1.07)
Proficiency-Writing	5.20 (1.20)

Materials and Procedure

All tasks were completed in a single day on a Mac computer using SuperLab experiment presentation software (Haxby et al., 1993). The shortened reading span task (Oswald et al., 2015) was completed on a PC computer using E-Prime software (Psychology Software Tools, Inc., 2016). The total procedure time ranged from 30 to 60 minutes depending on the English proficiency of the participant.

Language History Questionnaire

Language history and proficiency information was gathered using the Language History Questionnaire (LHQ) developed by the Brain, Language and Computation Lab at Pennsylvania State University (Li, Zhang, Tsai and Puls, 2014). Participants completed the questionnaire on a laptop and responses were downloaded for analysis at the end of the study.

English Vocabulary Test

The materials included the same 40 items used in experiment #1. Ten vocabulary item questions at the middle school learning level were added to make the test easier for this participant group. The new vocabulary items added for experiment #2 are the final ten items in Appendix D.

Reading Span Test

The materials and procedure for this task are identical to experiment #1.

Lexical Decision Test

The materials and procedure for this task are identical to experiment #1.

Masked Priming Test

The materials and procedure for this task are identical to experiment #1. Like the first experiment, participants were asked if they were able to read the word primes after completing the task.

Data Analysis

Participants were eliminated from the various tasks as follows: Four participants completed an older version of the vocabulary test and were excluded from the vocabulary score analysis. Four participants did not complete the language history questionnaire and I was unable to collect age of acquisition information from these subjects. These four participants were excluded from any ANOVA or mixed model test using either vocabulary or age of acquisition as a factor in the model. This means four participants were excluded from both the lexical decision and masked priming task when vocabulary group and age of acquisition was used as a factor in the models.

All participants except one earned at least 80% accuracy on the grammaticality judgments for the reading span test. The one participant who earned less than 80% accuracy on the task was excluded from the reading span data. Seven participants stated that they were able to read the word primes in the priming task and so their data was also excluded from the analysis on the priming task. One participant used their non-dominant hand for the masked priming task and their dominant hand for the lexical decision task and so their masked priming data was excluded from analysis. This left 37 participants included in the analysis when vocabulary and age of acquisition was a factor and 41 participants when it was not a factor.

No vocabulary words on the English vocabulary test earned less than 10% accuracy and therefore no words were excluded. In experiment #1, three highly frequent British English vocabulary words but low frequency American English words and four English target words that earned less than 80% accuracy were excluded from the data analysis on the lexical decision task. This was done in order to provide an effective baseline for the later experiments. In this population, the same seven words also earned less than 80% accuracy and were also excluded from the analysis. In both experiments #1 and #2 the same seven words were excluded from the English word lexical decision analysis. For this experiment, the accuracy rates for those words are: W-S- docile, *M*=.51, W+S+ cookery, *M*= .65, W-S+ hoover, W+S- barrow, *M*= .67, W-S+ mirage, *M*=.69, W+S+ bony, *M*=.71 and W+S+ duster, *M*= .73. This removed 8.75% of the total English word data.

The same three nonwords as the native adult dataset were removed so the words matched the native group analysis. The L2 group average for those words are: W+S+ peacher, *M*=.35; W+S+ proudy, *M*=.35, W+S+ clockage, *M*=.37. This removed 3.75% of the data.

English age of acquisition in listening, speaking, reading and writing, English vocabulary and reading span performance along with total years of English language were correlated with participant performance on the experimental tasks using a Pearson's Correlation test. English reading acquisition age (the age at which the participant began to read in English) was selected as the age of acquisition variable for the statistical models. This variable was chosen because it was the only age of acquisition domain that did not correlate with chronological age. I wanted to choose an English acquisition variable that would not overlap with chronological age in order to avoid differences due to age effects and not English exposure.

For the statistical analyses of the experimental tasks, reading age was divided into two groups based on the median age of English reading acquisition which was seven years old. Participants who began to read in English before age 7 were assigned to the early age of acquisition group and participants who began to read in English at 7 or older were assigned to the late age of acquisition group. In the ANOVA models, English reading age was entered as a between-subjects variable. In the mixed effects models, reading age was entered as a fixed factor.

Repeated measures ANOVAs and mixed model analyses were conducted for both the lexical decision and masked priming results. When the mixed model matched the ANOVA effects or matched and then added to the results, only mixed model results were reported in order to avoid redundancy. In one case, the ANOVA produced an effect that was not present in the mixed model and so both results were reported in this section. All repeated measures ANOVA tests used both subject and item analyses and then minF′ was calculated. ANOVA tests were computed using SPSS software (IBM Corp, 2017).

Mixed effect models were computed in R (R Core Team, 2017) using the package lme4 (Bates, Machler, Bolker & Walker, 2014). A maximal model approach was selected for both linear and logistic mixed effects models for this experiment. A description of the statistical models are included before the results for each task. The statistical output from all mixed effect models are included in Appendix G.

Results

English Vocabulary Test

Forty-five participants completed the vocabulary test during the experimental procedure and the average score was 51% accuracy with a standard deviation of 13%. Each vocabulary item had five answer options and chance was 20% accuracy. A one-sample t-test shows that the group performed significantly above chance, $t\ (44) = 16.34$, $p < .001$, 95% CI, .28 - .35. The average response time to choose a vocabulary definition was 7.40 seconds with a standard deviation of 2.90 seconds.

Reading Span Test

Thirty-nine participants completed the Reading Span test of the forty-nine participants total. The average partial score was 15.49 with a standard deviation of 7.80 points. This test was given last in the experimental procedure and like the first experiment, some participants ran out of time and so were unable to complete the task.

Lexical Decision Task

Data from 45 participants were included in the statistical analyses. All tests were conducted for English words and nonwords separately. Accuracy and response times were analyzed by applying a 2 (Base present/absent) X 2 (Suffix present/absent) X 2 (Reading Age Early/Late) ANOVA test using SPSS (IBM Corp, 2017). For the subject analysis, the first two variables are within factors and vocabulary group is a between factor. For the item analysis, all three variables are between factors. Subject (F1) and Word Item (F2) were analyzed as random variables and minF′ was then computed (Clark, 1973).

Incorrect responses were excluded from the response time analysis and any response slower than 2.5 standard deviations above the mean per subject were also excluded. This removed 5.23% of the English word responses and 25.54% of the Nonword responses. Table 6 shows the average accuracy and response times for the English and nonword categories separated by age of reading acquisition group. Because a considerable amount of data was excluded in the nonword response time analysis, I assume most of the effects will be relegated to the accuracy analysis.

Both accuracy and response times were analyzed using a mixed effect model in R using the package lme4 (Bates et al., 2014). For all mixed effects analyses for this task, the base, suffix and vocabulary group variables were entered as fixed effects. Subject and word item variables were entered as random intercepts and random slopes were selected for the subject random effect. Random slope parameters are specified before the results are discussed in each section. Effects were reported using the same procedure as the previous experiment with z-values and Wald's z-test for accuracy and t-values and Wald's chi-square test for response times.

English Word Accuracy

There was a main effect of Base for subjects but not items, F_1 (1, 43) = 7.20, $p < .05$, $\eta_p^2 = .14$, $F_2 < 1$ and there was no main effect of Suffix, $F_1 < 1$, $F_2 < 1$. There was no main effect of Reading Age group, $F_1 < 1$, $F_2 < 1$.

The interaction between the Base and Suffix variable was significant, minF′ (1, 178) = 4.07, $p < .05$, (F_1 (1, 43) = 13.95, $p < .01$, $\eta_p^2 = .25$, F_2 (1, 138) = 5.75, $p < .05$, $\eta_p^2 = .04$). For subjects, when a word did not feature a base word adding a suffix to the word improved accuracy (W-S-, *M*=.93, *SE*=.01, W-S+, *M*=.97, *SE*=.01), *F* (1, 44) = 8.12, $p < .01$, $\eta_p^2 = .16$. When a word featured a base word adding a suffix to the word decreased accuracy (W+S-, *M*=.98, *SE*=.01, W+S+, *M*=.95, *SE*=.01), *F* (1, 44) = 5.75, $p < .05$, $\eta_p^2 = .12$. This interaction is depicted in Figure 6.

Table 6. Adult non-native English Speaker means and standard deviation values for accuracy and response times on the lexical decision task.

	W+S+	W+S-	W-S+	W-S-
English Accuracy (%)				
Early Reading Age	95.40 (6.13)	98.17 (3.41)	97.19 (5.57)	93.91 (7.82)
Late Reading Age	94.65 (5.12)	96.89 (5.05)	95.99 (4.93)	92.95 (7.01)
English RT (ms)				
Early Reading Age	715.18 (109.31)	692.81 (136.21)	680.44 (113.70)	699.45 (113.03)
Late Reading Age	786.00 (139.10)	752.73 (117.63)	770.41 (114.41)	771.79 (96.63)
Nonword Accuracy (%)				
Early Reading Age	75.96 (20.22)	84.78 (16.20)	90.00 (12.34)	88.48 (9.35)
Late Reading Age	59.09 (28.33)	70.23 (25.05)	77.05 (25.06)	75.45 (22.30)
Nonword RT (ms)				
Early Reading Age	919.66 (177.06)	879.53 (156.33)	874.08 (167.36)	847.14 (196.55)
Late Reading Age	1026.16 (277.93)	1079.13 (280.76)	980.48 (203.20)	957.53 (187.06)

For items, the pattern was identical as the subject comparisons but neither simple comparison was significant. When a word did not feature a base word adding a suffix to the word improved accuracy (W-S-, *M*=.93, *SE*=.01, W-S+, *M*=.98, *SE*=.01), *F* (1, 70) = 3.10, *p* =.09. When a word featured a base word adding a suffix to the word decreased accuracy (W+S-, *M*=.96, *SE*=.01, W+S+, *M*=.95, *SE*=.01), *F* (1, 72) = 2.87, *p* =.09. There were no other significant interactions for the subject or item random effect analyses.

A mixed effect logistic regression was conducted using the binomial link function and a base by suffix interaction plus reading age group were entered as fixed effects. Subjects and word items were entered as random effects and were given random intercepts. I was unable to apply a maximal model approach for this model as no random slope terms produced a fitted model. This inflates the Type 1 error of the results.

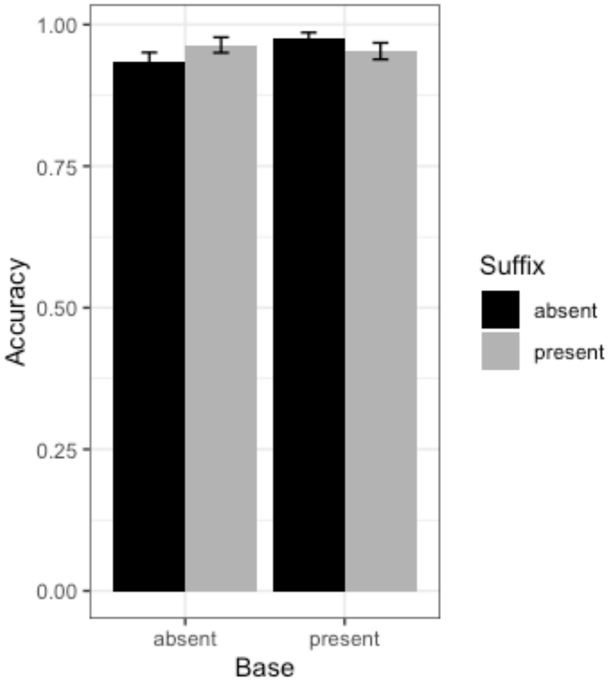


Figure 6. Average accuracy for base and suffix variables on English targets in non-native adult English speakers.

The model showed no significant effect for the variables base, *z*= 1.94, suffix, *z*= 1.33 or reading age group, *z*=-.95. There were also no significant interactions between any of the variables including base by suffix, *z*= -1.72. These results do not match the ANOVA results since a significant interaction between base and suffix was identified in the ANOVA but not the mixed effect logistic regression model.

Previous research has used English vocabulary proficiency as an individual difference factor in non-native English samples. I chose to use age of acquisition as a variable because I hypothesized that duration of exposure to English might be a better individual difference measure on these tasks than vocabulary knowledge. However, in order to better compare my study to previous research, I will also add vocabulary group to the mixed models for all dependent variables in this experiment. Like the native English adult group, vocabulary accuracy z-score was computed for each participant and participants were assigned to a high and low group based on whether their z-score was below or above 0.

For the accuracy model, a base, suffix and vocabulary group interaction was entered as a fixed effect and the random effect of subject was assigned a random slope of base. For this model the random effects were entered separately in order for the model to converge appropriately. The alternative model showed a main effect of vocabulary group, $z_{subject}= 4.41$, $p < .05$, $z_{word}= 2.88$, $p < .01$ in that the high vocabulary group earned .90 more accuracy points on average. There were no other significant main effects or interactions.

English Word Response Time

Results from response times are reported from the linear mixed effects model as they did not differ from the ANOVA tests. A base, suffix and reading age interaction was entered as a fixed effect while the subject and word item variables were entered as random effects. The subject random effect was given a random intercept and a random slope of a base by suffix interaction. The word random effect was left as a random intercept. The random slope was chosen based on the ANOVA model which showed this interaction was significant for subjects. This model did not converge and the random effect subjects was given a random slope by base because this is the most complex model that would converge.

The alternative model showed no main effect of suffix, $t = .77$ or base, $t = .81$. There was a main effect of reading age group which matched the pattern of the ANOVA. Participants in the older English reading age group were 74.10 milliseconds slower on average to respond to a target word, $t = -2.39$, $p < .05$. There were no significant interactions including the base and suffix interaction, $t = -1.50$.

When vocabulary group was entered into the model with the same parameters as the reading age model, there was a main effect of vocabulary group, $t = 3.01$, $p < .01$. Specifically, the high vocabulary group was 94.63 milliseconds faster to respond to target words on average than the low vocabulary group.

The results of the linear mixed effect model match the ANOVA in that none of the main effects or interactions are significant except for the effect of reading age and vocabulary group. The results show participants are not responding to target words differently based on the base and suffix manipulations in the target word.

Nonword Accuracy

Subjects and items random effects ANOVAs were computed followed by minF′ analyses. A mixed model logistic regression was then computed in R. The mixed model matches and expands on the results of the ANOVAs so only the mixed model results are given here to be as succinct as possible.

A mixed effect logistic regression was conducted using the binomial link function. A base, suffix and reading age group interaction was entered as a fixed effect and subjects and word items were entered as random effects and were given random intercepts. Random slopes were not assigned to this model as none of the random slope terms allowed the model to converge. This inflates the Type 1 error rate of the model.

The model showed that the variables base, $z=-.32$, suffix, $z=.97$ and reading age, $z=-1.92$ were not significant predictors of nonword accuracy. There was a significant interaction between the base and suffix variable as a predictor of nonword accuracy, $z=-2.20$, $p <.05$. There were no other interactions. A simple effects analysis for the base and suffix interaction was conducted by computing simple mixed models on each level of the base word variable. Suffix was entered as a fixed effect and was selected for the random slope for subjects.

When there was no base word present in the nonword, adding a suffix to the word did not predict target word accuracy, *z* = 1.22. When there was a base word present in the nonword, adding a suffix to the word decreased the likelihood of accurately responding to the target nonword by .85 points on average, *z*= -2.49, *p* < .05. This interaction is shown in Figure 7.

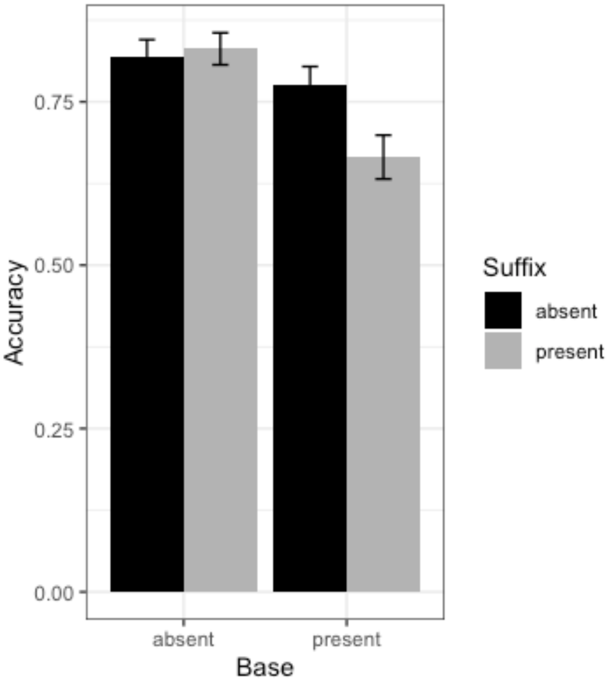


Figure 7. Average accuracy for base and suffix variables on nonword targets in non-native adult English speakers.

The vocabulary group variable was entered as a fixed factor with the same model parameters as the reading age model and vocabulary group was not a significant predictor of nonword accuracy, *z*=1.30 but the base variable was, *z*=-2.21, *p* <.05. Non-native speakers were .74 points less accurate on average for base present nonwords. There was also a significant interaction between base and vocabulary group, *z*=2.68, *p* <.01, suffix and vocabulary group, *z*=2.38, *p* <.05 and a three-way interaction between base, suffix and vocabulary group, *z*=-2.60, *p* <.01.

For the base variable, the low vocabulary group showed a greater accuracy cost to responding to a nonword with a base present, *z*=-3.64, *p* <.001 compared to the high vocabulary group, *z* =-2.70, *p* <.01. The low vocabulary group scored .94 points lower on base present nonwords compared to .79 points for the high vocabulary group. This interaction is depicted in Figure 8.

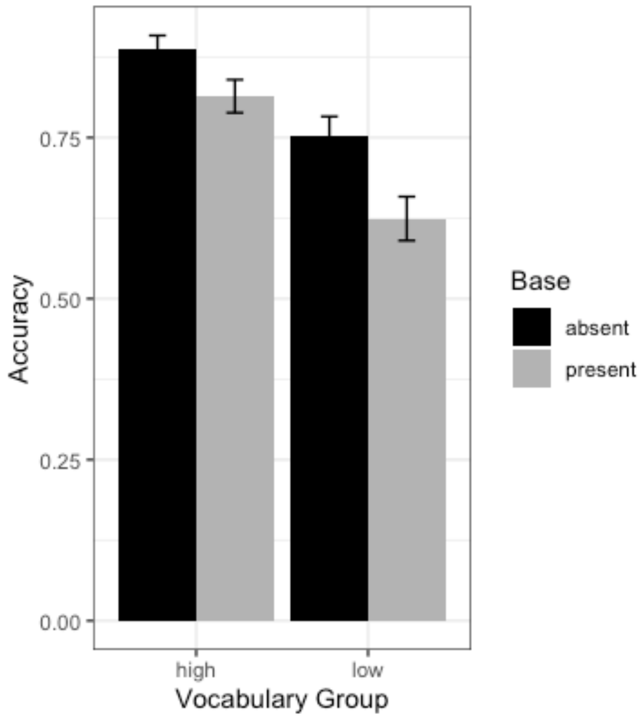


Figure 8. Average accuracy for the base variable on nonword targets separated by vocabulary group in non-native adult English speakers.

For the suffix variable, the low vocabulary group was significantly less accurate on suffix present nonwords, *z* =-1.96, *p* <.05 but the high vocabulary group was not *z* =-1.09. This interaction is depicted in Figure 9.

Finally, there was significant interaction between the base and suffix variable for the high vocabulary group, *z*=-2.73, *p* <.01 but not the low vocabulary group, *z*=-.96. When there was no base word in the nonword, adding a suffix did not affect nonword accuracy for the high vocabulary group, z=1.77. When there was a base word in the nonword, adding a suffix significantly decreased nonword accuracy for the high vocabulary group by .78 points on average, z=-2.11, p <.05. This interaction is depicted in Figure 10.

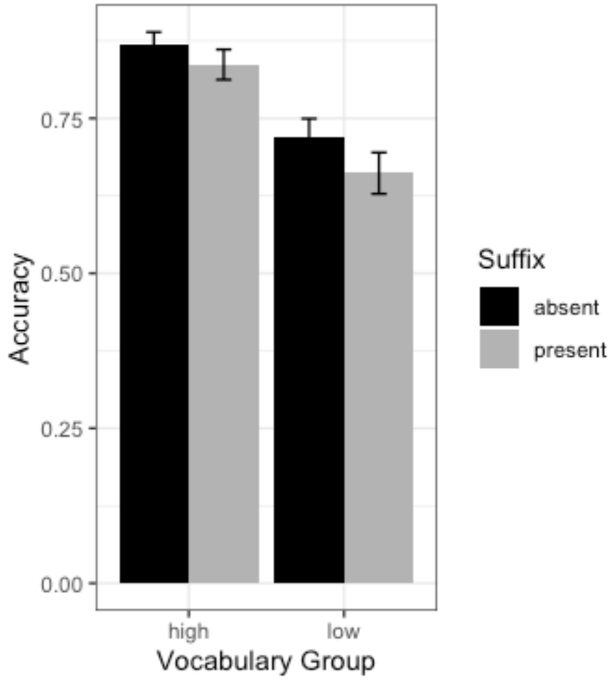


Figure 9. Average accuracy for the suffix variable on nonword targets separated by vocabulary group in non-native adult English speakers.

Nonword Response Time

Like the accuracy dependent variable, subjects and items random effects ANOVAs were computed followed by minF′ analyses. A linear mixed model was then computed in R. The mixed model matches and expands on the results of the ANOVAs so only the mixed model results are given in order to not be repetitive.

A linear mixed effects model with a base, suffix and reading age group interaction was entered as a fixed effect and subjects and word items were entered as random effects and were given random intercepts. Base by suffix interaction was entered as a random slope for the subject random effect but this model did not converge. Base was entered as the random slope for subjects as this was the most complex model that would converge.

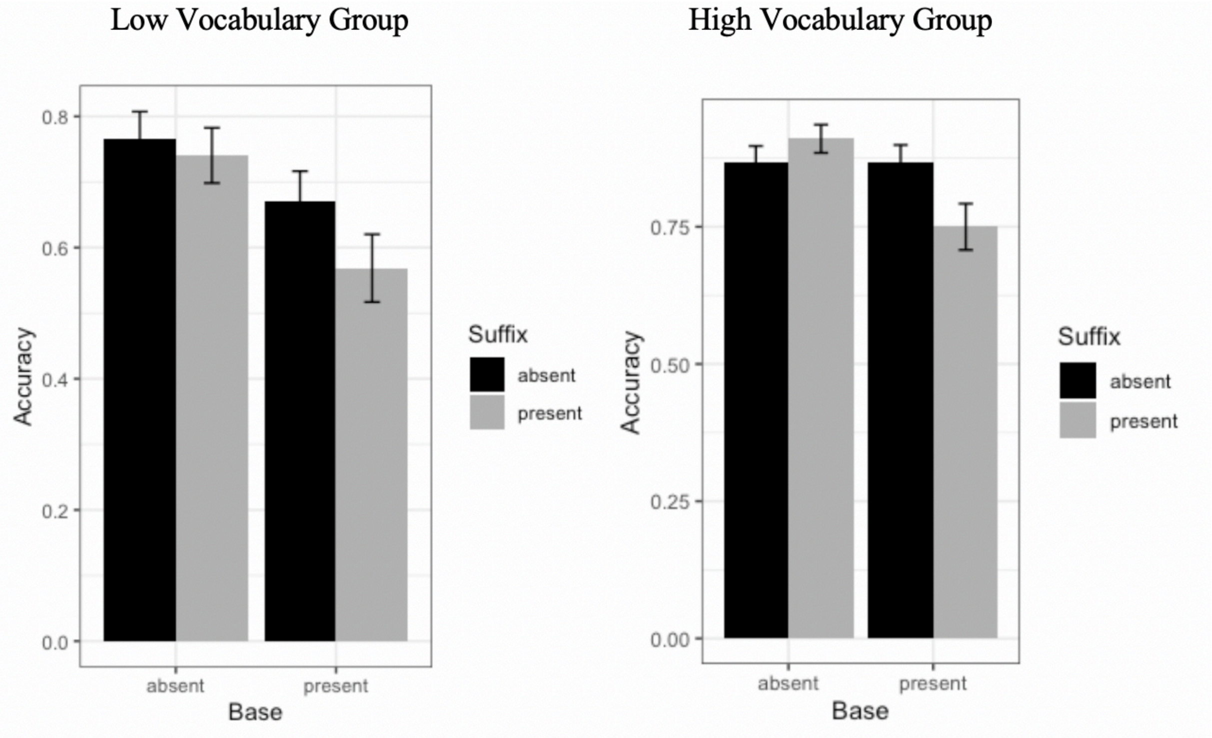


Figure 10. Average accuracy for the base and suffix variables on nonword targets separated by vocabulary group in non-native adult English speakers.

The model showed that the variables base, *t*=-1.87 and suffix, *t*=-.12 did not significantly predict nonword response times but reading age did, *t*=-3.16, *p* <.01. When a participant reported beginning to read English age 7 years or later, the average response time increased 164.64 milliseconds on average.

There were no significant interactions including the base and suffix variables, *t* = -.97 interaction. These results show that participants are responding to nonword targets based on the manipulations in the word and based on the duration of English reading exposure.

Next, vocabulary group replaced reading age as a fixed factor and was entered using the same model parameters indicated above, the results showed a main effect of vocabulary group, *t*=2.12, *p* <.05 and no interactions relating to that variable. Participants in the high vocabulary group were significantly faster to correctly respond to nonword targets compared to the low vocabulary group by 110.97 milliseconds on average.

Masked Priming Task

Data from 37 participants were analyzed for English word trials only. Accuracy was analyzed with a 2 (Relatedness) X 3 (Condition) X 2 (Reading Age Early/Late) mixed factorial ANOVA. The first two variables are within factors and vocabulary group is a between factor for the subject analysis whereas relatedness is a within factor and condition and reading age group are between factors for the item analyses. Subject (F1) and Item (F2) variables were computed as random effects variables and minF′ was then computed.

Incorrect responses were excluded from the response time analysis. Response times faster than 150 milliseconds and slower than 2.5 standard deviations above the mean were also excluded. This removed 3.54% of the responses. Table 7 shows the average response times and error rate for each prime and relatedness category.

Response times were first analyzed by applying a 2 (Relatedness) X 3 (Condition) X 2 (Reading Age Early/Late) mixed factorial ANOVA. As stated above, the first two variables are within factors and reading age group is a between factor for the subject analysis and all three variables are between factors for the items analysis. Subject (F1) and Item (F2) variables were computed as random effects variables and minF′ was then computed.

For mixed effect model analyses, a maximal model approach was applied for both the accuracy and response time models. Relatedness, condition and reading age group variables were entered as fixed effects. Subject and word item variables were entered as random effects and given random intercepts. Random slopes were selected first based on the study design and then removed in order to produce a model that converged. Test values and distributions used for reporting effects are identical to the previous sections.

Table 7. Adult non-native speaker means and standard deviation values for accuracy and response times on the masked priming task.

	Form (castle/cast)	Opaque (corner/corn)	Transparent (hunter/hunt)
Related Prime Accuracy (%)			
Early Reading Age	98.26 (3.88)	95.65 (6.62)	99.13 (2.88)
Late Reading Age	95.24 (8.14)	94.76 (6.80)	96.67 (5.77)
Related Prime RT (ms)			
Early Reading Age	1245.67 (79.04)	1261.01 (79.85)	1216.36 (61.22)
Late Reading Age	1310.58 (109.40)	1324.94 (109.53)	1292.71 (93.20)
Unrelated Prime Accuracy (%)			
Early Reading Age	97.39 (6.89)	96.52 (7.14)	97.83 (5.18)
Late Reading Age	93.81 (6.69)	97.62 (4.36)	96.19 (5.90)
Unrelated Prime RT (ms)			
Early Reading Age	1252.31 (78.59)	1261.91 (66.71)	1243.41 (67.35)
Late Reading Age	1352.48 (93.10)	1329.08 (87.66)	1342.70 (88.27)

Accuracy

Subjects and items random effects ANOVAs were computed followed by minF′ analyses. A mixed model logistic regression was then computed in R. The mixed model matches and expands on the results of the ANOVAs so only the mixed model results are given below.

A mixed effect logistic regression was conducted using the binomial link function. A relatedness by condition interaction plus reading age group factor were entered as fixed effects and subjects and word items were entered as random effects and were given random intercepts. Relatedness by condition interaction was entered as a random slope for the subject random effect and relatedness was added as a random slope for the word random effect but this model did not converge. Relatedness was then removed from the subject random slope so that condition was assigned as a random slope for subjects and relatedness for word random effects. This model also did not pass the convergence test. Relatedness was then removed as a random slope from the word random effects term and this model still did not converge. Relatedness was added again as a random slope for word and condition was removed from the subject random effect. This model did converge.

The model results showed no significant effects of relatedness, $z= .95$, condition, $z=1.95$ and no interaction between the two, $z=-1.05$. The reading age variable did significantly predict target accuracy for the masked priming task, $z=-2.45$, $p < .05$. When a participant reported beginning to read English age 7 or later, their average accuracy score decreased .69 points on average.

Vocabulary group was entered in place of reading age with the same model parameters as above. Vocabulary group was a significant predictor of target accuracy with high vocabulary group participants earning .83 more points on average compared to the low vocabulary group, $z=2.98$, $p < .01$. No interaction terms with this variable were significant.

Response Time

Subjects and items random effects ANOVAs were computed followed by minF′ analyses. A linear mixed model was then computed in R. The mixed model matches and expands on the results of the ANOVAs so only the mixed model results are given.

A linear mixed effect model was entered with relatedness, condition and reading age group as fixed effects. Subjects and word items were entered as random effects and were given random intercepts. Relatedness by condition interaction was entered as a random slope for the subject random effect and relatedness was added as a random slope for the word random effect but this model did not converge. Relatedness was then removed from the subject random slope so that condition was assigned as a random slope for subjects and relatedness for word random effects. This model did not pass the convergence test and so relatedness was removed from the word random effect and this model did pass the convergence test.

The model showed the variables relatedness, t=4.03, p <.001, condition, t=2.08, p <.05 and reading age, t=-3.13, p <.01 were all significant predictors of the speed at which a participant correctly responded to a target word. For the relatedness variable, when a target's prime word was related to the target, the average response time decreased 20.40 milliseconds. For the reading age group variable, when a participant reported beginning to read English age 7 years or later, the average response time increased 79.29 milliseconds. There were no significant interactions including the relatedness and condition variable interaction, t = 1.06.

For the condition variable, a Tukey HSD test was conducted on the factor. Although the factor was significant in the linear mixed effect model, the F test was not, F (2) =2.61, p=.07. Even though there was not a significant interaction between the relatedness and condition variables, I entered relatedness as a fixed factor at each condition level to look for significant priming effects. For these simple models, relatedness was entered as a random slope for both subject and word random effects and these models converged appropriately. The form condition showed that the related condition was 22.12 milliseconds faster than the unrelated condition which shows a positive priming pattern that was not significant, t = -1.51. The opaque condition showed minimal positive priming with a 2.54 millisecond advantage in the related prime condition, t = -.20. The transparent condition showed a significant positive priming effect with participants responding 37.10 milliseconds faster in the related condition on average, t=-2.70, p <.01.

These results suggest non-native adult English speakers are mentally decomposing complex prime words in the and transparent condition but not in the form or opaque condition.

This is an interesting pattern that was identified as well in Li et al. (2017) and will be discussed at the end of this chapter and in the general discussion.

Like the previous sections, I also entered vocabulary group as a fixed factor using the same model parameters as above except, I was able to add relatedness as a random slope to both random effect terms and still have the model converge. The results show vocabulary group was a significant predictor of target response time, t=2.43, p <.05 but there were no significant interactions with this predictor. High vocabulary group participants were 63.34 milliseconds faster to respond to target words compared to the low vocabulary group.

Vocabulary and Age of Acquisition Correlations

Vocabulary test accuracy, reading span partial score and English age of acquisition and age of reading acquisition was correlated with lexical decision and masked priming accuracy and response times using a Pearson's correlation test. Reading span did not correlate with any variables aside from age of acquisition and so is therefore not included in the below correlation tables. English vocabulary accuracy correlated with lexical decision accuracy but did not relate with response times for either task. Neither age of acquisition or vocabulary performance correlated with masked priming accuracy. This is likely because accuracy on the masked priming task was already high.

Table 8 shows that English Vocabulary accuracy has the most consistent correlations with English target accuracy. The higher the participant scored on the English vocabulary test, the higher accuracy score they received on three of the four English target conditions. The conditions where there are correlations each have one English morpheme present, either base or suffix form. Vocabulary knowledge was beneficial for these conditions because word knowledge correlates with morpheme knowledge and it is likely the participant was using their vocabulary resources for these conditions. Age of reading acquisition negatively correlates with the W+S- English accuracy condition. The earlier a participant began reading English the higher they scored on the W+S- English accuracy condition. This may be another example of how the base form of a word is valuable to English word processing and earlier English readers in the L2 group maybe practiced in base word processing giving them an advantage in this condition.

Table 8. Adult non-native speaker correlation between lexical decision accuracy and age of English reading acquisition and English vocabulary accuracy.

Lexical Decision Accuracy	Age of English Reading Acquisition	English Vocabulary Accuracy
W-S+ English	$r = -.214$, $p = .16$	$r = .489$***, $p < .001$
W+S- English	$r = -.311$*, $p < .05$	$r = .339$*, $p < .05$
W+S+ English	$r = -.154$, $p = .31$	$r = .311$*, $p < .05$

General age of acquisition positively correlates with all English response time conditions and one nonword condition. This may be because accuracy was low in the nonword condition and so many trials were removed for analysis. It also could be because the conditions were difficult for all participants and so no individual differences are detected. Age of reading acquisition positively correlates with English targets that do not feature a base word and the W-S+ nonword condition. The earlier participants began to read English the faster their response times in these conditions. These relationships are shown in Table 9. Earlier reading experience may give participants a response time advantage in the base absent English target condition because they have learned a more efficient lexical access strategy for English words without base words compared to the late English readers.

Age of acquisition positively correlates with response times on all masked priming conditions. The earlier a participant acquired English the faster their response times on all

English target words. Accuracy was very low in that condition overall. Age of reading acquisition positively correlates with all conditions except the related form condition. This may be because the related form condition was difficult or more effortful to process for all L2 adult participants.

Table 9. Adult non-native speaker correlation between lexical decision response time and age of acquisition and age of English reading acquisition.

Lexical Decision Response Times	Age of English Acquisition	Age of English Reading Acquisition
W-S- English	$r = .353^*, p < .05$	$r = .338^*, p < .05$
W-S+ English	$r = .455^{**}, p < .01$	$r = .475^{**}, p < .01$
W+S- English	$r = .314^*, p < .05$	$r = .246, p = .10$
W+S+ English	$r = .330^*, p < .05$	$r = .276, p = .07$
W+S- Nonword	$r = .377^{**}, p < .01$	$r = .314^*, p < .05$

Table 10. Adult non-native speaker correlation between masked priming task response times and age of acquisition and age of English reading acquisition.

Masked Priming Response Times	Age of English Acquisition	Age of English Reading Acquisition
Unrelated Form Prime	$r = .477^{**}, p < .01$	$r = .406^{**}, p < .01$
Unrelated Opaque Prime	$r = .463^{**}, p < .01$	$r = .349^*, p < .05$
Unrelated Transparent Prime	$r = .558^{***}, p < .001$	$r = .468^{**}, p < .01$
Related Form Prime	$r = .313^*, p < .05$	$r = .258, p = .09$
Related Opaque Prime	$r = .384^{**}, p < .01$	$r = .328^*, p < .05$
Related Transparent Prime	$r = .517^{***}, p < .001$	$r = .399^{**}, p < .01$

These relationships are shown in Table 10. Language background and vocabulary test accuracy did not correlate with accuracy or response time priming effects in any of the conditions. Language and demographic background did not correlate with any of the priming effects in each condition.

Discussion

Adult non-native speakers of English completed an English vocabulary test, the shortened version of the reading span test, a lexical decision task and a masked priming task. My hypotheses before the experiment were that adult non-native speakers of English would be sensitive to word structure changes in the lexical decision task as reflected in their accuracy and response times. This would show they are mentally decomposing complex words during word processing. Second, adult non-native speakers would not engage in obligatory decomposition in the masked priming task and would most likely show positive priming in each of the priming conditions. I also believed that age of acquisition would influence task performance with participants who acquired English earlier in life, showing priming effects similar to native speakers. In this experiment, I used age of English reading acquisition because this was the only age of acquisition domain that did not correlate with chronological age.

My hypotheses for the lexical decision task were supported and non-native speakers did perform differently when base and suffix changes were made to the target word, particularly for nonword targets. My hypotheses for the masked priming task were partially supported, as there is some evidence that participants are behaving differently by condition. Additionally, participants in the early age of reading acquisition group (before age 7) were more accurate and faster to correct respond in some instances but there were no interactions, so these effects are not specific to any manipulations in the experimental tasks. Finally, like the native adult group, reading span performance did not correlate with any dependent measure.

Lexical Decision Task

Unlike the native speaking group, previous work has been done using this task with French-English bilingual adults (Casalis et al., 2015b). For English target words, accuracy showed that having a suffix present in a target word with a base word (W+S+) led to lower accuracy than not having a suffix present in a word with a base word (W+S-). When a word did not have a base word, adding a suffix to the word improved accuracy (W-S-, W-S+). This result was supported by the significant interaction in the ANOVA but the interaction was not present in the mixed model. No other effects were present in either statistical model. It is unclear why there was a significant interaction in the ANOVA test and no the mixed model. One possibility is that when considering the factor coefficients all together, the variability described by the interaction becomes less meaningful. It is not because the mixed model is eliminating possible type 1 error because no random slopes were assigned to the English accuracy model as it did not produce a converging model.

Casalis et al. (2015b) did not find the interaction effect in their study. Their non-native participants showed better accuracy for target word conditions that featured a base word or a suffix morpheme. This pattern makes sense as seeing a familiar English structure (base words and suffixes are replicable in other words) would lead to a higher probability of correctly accepting the target word as an English word. They also found that participants with low vocabulary proficiency had higher accuracy for words with a base word (Base and Vocabulary Group interaction) compared to high vocabulary proficiency participants who did not show accuracy differences based on the base word variable. The present experiment used age of English reading acquisition as a fixed factor in the models not vocabulary group, but I did enter vocabulary group into the mixed model to look for significant effects. Neither the age of reading acquisition nor vocabulary group factors influenced task performance for English words in this group.

The Casalis et al. (2015b) study was most likely able to identify accuracy differences based on vocabulary group because none of their participants were early bilinguals and their low proficiency group had not been to an English-speaking country for longer than a week. My population had an average age of acquisition of 6 years and all had spent at least 6 months in the United States (one semester).

There were no significant results in the response time data for English words beyond a faster response time overall for the early reading age group. The Casalis et al. (2015b) study also found no significant effects for the English response time data. This means that while the changes to the structure of the English target word influence participants ability to correctly respond, it does not affect speed of processing once the participant recognizes it as an English word. In my study, morphological decomposition may not be taking place for English target words as words with and without English morphemes are responded to at the same speed and at the same accuracy rate. Another explanation for why there were no response time effects is that the length of target presentation is not long enough to capture processing differences based on condition. However, this is unlikely since targets were shown for 5000 milliseconds and the average response time for every condition was under 800 milliseconds. I also believe my sample size was adequate for the task.

A morphological decomposition strategy is being used by this group for nonwords. In the accuracy data, there was a significant interaction between the base and suffix variables in that the most English-like (W+S+) condition showed the lowest accuracy. These results match identically what Casalis et al. (2015b) found.

The Casalis study found no interactions with the vocabulary group variable for nonword targets. The vocabulary group accuracy analysis in the present experiment was able to capture more details in how the nonword structure influenced the participants ability to correctly identify a nonword. For example, adding a base word or a suffix to the nonword decreased accuracy for low vocabulary participants but not for high vocabulary participants. This matches what Casalis found in her English word results with a base and vocabulary proficiency interaction for accuracy. In that case, adding a base word improved English word accuracy for the low vocabulary group. Combined with my results, this might suggest that when learning a new complex English word, having a base word present in the word helps the learner acquire it. A three-way interaction in my study shows the high vocabulary group is performing more like native English speakers in that they are not showing significantly worse accuracy for nonwords until the most English-like condition, the W+S+ condition.

Casalis et al. (2015b) did have a response time interaction between base and suffix that matched the pattern of the accuracy data in that the W+S+ condition had the slowest response times for nonwords. This experiment found no interaction in the nonword response time data. This is most likely because the low accuracy rates for nonwords overall led to over a quarter of the data being excluded for the response time analysis. This decreases the amount of useable data for the mixed model and decreases the chances of finding unique variance akin to a base and suffix interaction.

Masked Priming Task

There were no accuracy effects for this task besides early English readers and high vocabulary participants earning more points than their counterparts. This is most likely because accuracy rates were high overall and so there was not enough variance in the dataset to capture differences based on condition.

The response time results of the masked priming task partially supported my hypothesis for non-native English speakers. I initially stated that non-native speakers would show positive priming effects in all three conditions, transparent, opaque and form because they would be identifying relationships between words based on other word charateristics besides morphology like semantics and orthography. The non-native speakers were faster at responding to targets with related primes compared to unrelated primes. There was no significant interaction between relatedness and condition but there was significant positive priming in the transparent condition (hunter-HUNT). Non-native participants are clearly identifying the structural and semantic similarities between the prime and target word resulting in a priming effect.

There was also non-significant positive priming in the form condition (castle-CAST) but not the opaque condition (corner-CORN). This means that participants saw castle and cast as related words, but corner and corn were not related. This pattern does not match what the native speakers displayed in their priming effects. It is also a pattern that does not make sense at first pass. If, as a non-native speaker, the reader recognizes that corner and corn are not related, what makes castle and cast related? They are structurally similar in that both pair examples share the same orthographic base but the opaque prime has a suffix and the form prime does not.

One study that makes this pattern seem less strange is Li et al. (2017) on which the study materials are based. In their study, they gave the same masked priming task to three participant groups. A native English-speaking group, a high proficiency Chinese-English bilingual group and a low proficiency Chinese-English bilingual group. The native speakers matched the Rastle et al. (2004) pattern of positive priming in the transparent and opaque condition and no priming in the form condition. The high proficiency speakers showed positive priming in all three conditions, but the pattern matched the native speakers in that priming was strongest in the transparent condition and weakest in the form condition. For the low proficiency speakers, they showed the same priming pattern as my study. Positive priming in both the form and transparent condition and lower priming in the opaque condition.

Additionally, the Li et al. (2017) study had 40 participants in the high proficiency group and 60 participants in the low proficiency group while my study had approximately 20 participants in each vocabulary and age of acquisition group. This may explain why there were no individual difference effects on my masked priming task. Li et al. (2017) also had a more homogenous sample population with each participant speaking the same native language but, in my sample, a wide range of native languages were represented. This is another explanation for why it was easier for Li et al. (2017) to identify individual differences in masked priming effects.

In my study, the opaque condition has a 3 millisecond (ms) priming effect with 37 ms priming in the transparent condition and 22 ms priming in the form condition. These priming results suggest the L2 adult speakers are potentially using different word processing strategies. For the transparent condition, they are using a morphological decomposition strategy that is benefitting their lexical search resulting in faster response times in the related condition. For the opaque condition, they are using whole word access as the related condition is equivalent to the unrelated condition. For the form condition, they are matching the orthographic similarity between the two words which is also benefitting their lexical search in the related condition. They cannot use morphological decomposition for this condition because there is no suffix to decompose.

It is also possible that a single strategy is being used that matches like orthography between words but semantic relatedness in the transparent condition facilitates response time speeds and unrelated semantics in the opaque condition produces inhibits response times making the related opaque condition appear identical to the unrelated opaque condition.

In Li et al. (2017), the low proficiency group has a 30 ms opaque priming effect with 50 millisecond priming effects in both the form and transparent condition. Where the strength of the priming effects are different in the Li et al. (2017) study compared to mine, the pattern is the same. Both the Li et al. (2017) study and my study have small priming effects in the opaque condition and larger priming effects in the form and transparent condition. It is important to note I did not find a significant form priming effect in my study but Li et al. (2017) did.

However, my participant group is not a low proficiency group like the Li et al. (2017) study. All of the participants are currently enrolled in an undergraduate program in the United States and currently taking all classes in English. The average age of acquisition was 6.5 years which means that even after an average duration of 16 years of English exposure, their morphological decomposition strategy is non-native like. This suggests that strategies used to identify relationships between words based on orthography and morphology may be a result of some other factor besides English proficiency. For example, Brooks, Kwoka & Kempe (2017) found that adult English speakers learned Russian morphology better than other participants if they also performed well on a pattern recognition test. This skill might be a better evaluator of masked priming differences than English vocabulary and duration of English reading exposure.

It is also possible non-native speakers of English rely on lexical access more than morphological decomposition as a word processing strategy and lexical access may be a strategy that the participants are using in the related opaque condition as it is the same speed as the unrelated condition. In a review of L2 morphology studies, Clahsen, Felser, Neubauer, Sato and Silva (2010) found that non-native speakers of a language do not rely on morphological processing to the same degree as native speakers and that whole word access is a more common strategy. This may be because morphological decomposition requires the reader to have learned the rules of how English morphology works which depending on the quality of their instruction and exposure, may not be the case in my participant sample.

This explanation is still not entirely satisfying. A whole word access strategy is typically associated with low English proficiency (Liang & Chen, 2014) and as stated previously, this participant group is highly proficient.

Another explanation is that for words like *corner*, non-native speakers are using a morphological decomposition strategy but because -er does not give meaning to the word *corner* in the same way -er gives meaning to the word *hunter*, interference occurs which slows processing. More research is needed to clarify what individual difference measures influence when morphological decomposition is and is not used in non-native speakers of English.

Chapter 5. Experiment 3. Elementary-Age Children

For the third and final experiment, elementary-age children in grades three through six completed the same two experimental tasks as the adult groups. They also completed a receptive vocabulary test which was correlated with task performance.

My research questions for experiment #3 were whether elementary age children, who are in the process of mastering English vocabulary, are sensitive to word structure changes in English words and nonwords. Following previous literature, I hypothesized children would be sensitive to word structure in the lexical decision task and use both base words and suffix structures to influence their lexical decisions.

Second, I was interested in whether the priming task would successfully produce clear priming effects in the three conditions: form, opaque and transparent prime-target pairs. Masked priming techniques are not always successful in this age group and have shown negative priming effects in one study (Beyersmann et al., 2012). I want to identify interpretable patterns in this group's response time data, whatever they may be. My hypothesis is that I would see priming effects in all conditions because children are still learning to morphologically decompose complex words during visual recognition. I also believed older age children (grades 5 and 6) would produce priming effects closer to adult native speakers, as they are the most likely to use decomposition strategies consistently in this participant group. This would be shown in the form of priming effects in the transparent and opaque conditions but not the form condition. I was not able to test this hypothesis as there was not a large enough sample size to include age or grade as a factor in the statistical models.

Additionally, participants who score in a higher vocabulary percentile compared to the rest of the sample, will be more likely to show a positive correlation in accuracy and a negative correlation with response times in the lexical decision and masked priming tasks. If participant behavior follows previous research, higher vocabulary participants should be more accurate and faster to correctly respond to less English-like words (W-S-) and more English-like nonwords (W+S+) in the lexical decision task. Higher vocabulary percentile scores may negatively correlate with form priming in that the higher the vocabulary score, the less the reader is going to see castle and cast as a related pair based on their knowledge of English word structure.

Method

Participants

Thirty children in grades three through six were recruited from Baton Rouge area elementary schools, summer camps and by word of mouth in the spring and summer months towards the end of the school year. Parent permission and child assent was obtained before testing began and school administrator permission was obtained when recruiting was done using school email accounts. Children were told they would be helping researchers understand how children learn words and that they would be playing word games on a laptop. Children were informed that they could stop at any time and that if they chose to do so, there would be no consequence. An adult witness observed child assent and added their signature to the child assent form underneath the child's signature.

All of the children were exposed to English from birth except for one participant who was excluded from all data analysis which left twenty-nine participants. No other participants were excluded. Sixteen of the included participants were females. The average participant was in 4th grade (*M*=4.10, *SD*=1.01) and was 10 years old (*M*=10.38, *SD*=1.08). The average participant was in the 83rd percentile for receptive vocabulary (*M*=83.06, *SD*=17.15).

Before participation, each child's parents completed a background questionnaire about their child's language abilities and habits. Nineteen of the parents stated that their child's vocabulary was above the rest of the class and fourteen of the parents stated that their child's reading ability was above the rest of the class. The rest of the parents said their children's vocabulary and reading was at the same level as the rest of the class. No parent felt their child's vocabulary or reading proficiency was below their classroom peers.

Eleven of the twenty-nine participants had some second language exposure (French n=7, Spanish n=9, Igbo n=1, Chinese n=1, Japanese n=1, Latin n=1) however the average percentage of daily English use was almost 100% (*M*=99.73, *SD*=.63). The average participant spent 5 hours reading a week (*M*=5.46, *SD*=4.19) and almost 30 minutes a week were spent being read to by a parent or teacher (*M*=.48, *SD*=.88).

Six children had received speech therapy at some point in their education, one child had received assistance for an audio processing disorder and one child had received an ADD diagnosis. No child had received assistance for a reading problem however two children had received reading tutoring. Three of the twenty-nine children received math tutoring. A copy of the language exposure and habits questionnaire is included in Appendix C.

Materials and Procedure

The experimental tests were given on a laptop using SuperLab software (Haxby et al., 1993). Vocabulary test scoring was completed using paper and pencil. Testing was done either at a laboratory on campus or at the child's aftercare or camp location in a quiet classroom. The procedure took approximately 35 minutes to complete depending on how far the participant got on the PPVT-4 test. Children who got far on the PPVT-4 test took longer to complete the procedure. Children were able to complete both the lexical decision and masked priming task in about twelve minutes.

Vocabulary Test

All participants except one completed a receptive vocabulary test. One participant asked to end the vocabulary test towards the end as the words became more difficult and so that participants data was excluded from analysis. Each child completed the Peabody Picture Vocabulary Test – version 4, form A (PPVT-4). Their performance on this test was also correlated with performance on the experimental measures. I did not add the reading span measure for this participant group because I wanted to keep the experimental procedure as short as possible.

Lexical Decision Task

The words and procedure used were identical to experiments 1 and 2.

Mask Priming Task

The words and procedure used were identical to experiments #1 and 2. This task was given before the lexical decision task in the experimental procedure. To orient the children to the task, participants were given instructions that a space alien had come to earth and we needed to teach the alien English words. However, a mistake had been made and silly, made up words had gotten mixed in with the English words in the computer program. The researcher needed the participants to help decide which words were good to teach the alien and required a “yes” answer and which were not good to teach the alien and required a “no” answer.

Participants were able to practice making lexical decisions with index cards first. The researcher held up an index card with eight hashmarks on one side and then turned over the card to show a “target” word. The participant verbally responded yes or no as a lexical decision. The hashmarks were used to illustrate what the child would see on the masked priming task (the forward mask). After the researcher felt comfortable they understood the task, the participant was able to start the practice portion of the task on the laptop. All participants understood the task instructions after completing the practice set. The index cards were not used for the lexical decision task because all children understood the task instructions after completing the masked priming task.

Participants were not asked if they could read the word prime after completing this task because during pilot testing with elementary participants, the children did not understand the question and they became confused and uncomfortable. No participant voluntary reported being able to see anything flash on the screen.

Data Analysis

Vocabulary test performance, age and grade was correlated with participant performance on the experimental tasks using a Pearson Correlation test. One participant was in the 37th percentile for vocabulary but their accuracy scores for English target words on both the lexical decision (M=.86) and masked priming task (M= .85) were not different from the average accuracy on those tasks ($M_{lexicaldecision}$= .90) and ($M_{maskedpriming}$= .90) and so they were not excluded from any analysis. Because of the small sample size, neither vocabulary score, age nor grade was not included as a factor in any of the complex statistical models for the experimental tasks. Additionally, the groups were not evenly distributed as there were 11 participants in the upper elementary group (5 and 6th grade) and 18 participants from the lower age group (3 and 4th grade).

In experiment 1, three highly frequent British English vocabulary words but low frequency American English words and four English target words that earned less than 80%

accuracy were excluded from the data analysis on the lexical decision task. This was done in order to provide an effective baseline for the later experiments. Like experiment #2, these same words were excluded for this data analysis. For this experiment, the accuracy rates for those words are: W-S+ docile, *M*=.19, W-S- quarrel, *M*= .33, W-S+ mirage, *M*=.38, W-S+ hoover, *M*= .48, W-S+ leisure, *M*=.48, W+S+ cookery, *M*=.52 and W+S- barrow, *M*= .71. This removed 8.75% of the total English word data.

The same three nonwords as the previous two experiments were excluded from the lexical decision task (W+S+ proudy, *M*=.19, W+S+ peacher, *M*=.38, W+S+ clockage, *M*=.38). This removed 3.75% of the data.

Repeated measures ANOVAs and mixed model analyses were conducted for both the lexical decision and masked priming results. When the mixed model matched the ANOVA effects or matched and then added to the results, only mixed model results were reported in order to not be redundant. In one case, the ANOVA produced an effect that was not present in the mixed model and so both results were reported in this section. All repeated measures ANOVA tests used both subject and item analyses and then minF′ was calculated. ANOVA tests were computed using SPSS software (IBM Corp, 2017).

Mixed effect models were computed in R (R Core Team, 2017) using the package lme4 (Bates, Machler, Bolker & Walker, 2014). A maximal model approach was selected for both linear and logistic mixed effects models for this experiment. A description of the statistical models are included before the results are reported for each task. The statistical output from all mixed effect models are included in Appendix G.

Results

Vocabulary Test

Twenty-eight participants completed the standardized vocabulary test during the experimental procedure and because vocabulary ability increases significantly by age and grade level, participant performance was evaluated based on percentile score. The average percentile score was 83 (*M*=83.06, *SD*=17.15) and the range was 37-99.9.

Lexical Decision Task

Data from 29 participants was included in the statistical analyses. All tests were conducted for English words and nonwords separately. Accuracy and response times were analyzed by applying a 2 (Base present/absent) X 2 (Suffix present/absent) ANOVA test using SPSS (IBM Corp, 2017). For the subject analysis, the both variables are within factors. For the item analysis, both variables are between factors. Subject (F1) and Word Item (F2) were analyzed as random variables and minF′ was then computed (Clark, 1973).

Incorrect responses were excluded from the response time analysis and any response slower than 2.5 standard deviations above the mean per subject were also excluded. This removed 6.90% of the English word responses and 15.66% of the Nonword responses. Table 11 shows the average accuracy and response times for the English and nonword categories. I did not have a large enough sample size to include grade level as a factor however I thought it would be informative to include the means of these groups in the tables for comparison.

Both accuracy and response times were analyzed using a mixed effect model in R using the package lme4 (Bates et al., 2014). For all lexical decision task mixed effects analyses, the base and suffix variables were entered as fixed effects. Subject and word item variables were entered as random intercepts and random slopes were selected for the subject random effect.

Random slope parameters are specified before the results are discussed in each section. The test distributions used for reporting are identical to the previous two experiments.

English Word Accuracy

The repeated measures ANOVA tests matched the results of the mixed model and so only the mixed effects model results are reported. A mixed effect logistic regression was conducted using the binomial link function to test accuracy using a more powerful statistic application. A base and suffix interaction was entered as a fixed effect and subjects and word items were entered as random effects and were given random intercepts. Maximal model parameters did not produce a fitted model and so Type 1 error is inflated for this model.

Table 11. Elementary age native English speaker means and standard deviation values for accuracy and response times on the lexical decision task.

	W+S+	W+S-	W-S+	W-S-
English Accuracy (%)	92.15 (10.34)	93.65 (7.75)	91.81 (10.84)	91.38 (8.85)
English RT (ms)	963.02 (145.42)	961.27 (136.13)	984.08 (149.35)	980.20 (149.46)
Nonword Accuracy (%)	66.94 (18.43)	78.45 (15.07)	84.83 (16.77)	83.97 (14.90)
Nonword RT (ms)	1240.50 (233.75)	1167.00 (183.97)	1111.42 (197.11)	1083.55 (175.75)

The model showed no significant effect for the variables base, z= .24, suffix and z= .02. There was also no significant interaction, z= -.28. This shows that the children were not using visual word processing strategies based on the morphological structure of the English word. Since the average vocabulary percentile score was in the 83rd percentile, it is also possible that these words are processed so quickly any strategies used to make word decisions look the same across conditions.

English Word Response Time

Results from response times are reported from the linear mixed effects model as they did not differ from the ANOVA tests. A base and suffix interaction was entered as a fixed effect while the subject and word item variables were entered as random effects. The subject random effect was given a random slope by base variable because this was this variable that showed the most difference between subjects in the means and this was the most complex model that converged appropriately.

The alternative model showed no main effect of base, $t = 1.48$ or suffix, $t = .14$. There was also no significant interaction between the base and suffix variables, $t = -.40$. These results show that the strategies used to process the words in each condition are equally efficient or inefficient.

Nonword Accuracy

Subjects and items random effects ANOVAs were computed followed by minF′ analyses. A mixed model logistic regression was then computed in R. The mixed model matches the results of the ANOVAs so only the mixed model results are given here to be as succinct as possible.

A mixed effect logistic regression was conducted using the binomial link function. A base and suffix interaction was entered as a fixed effect and subjects and word items were entered as random effects and were given random intercepts. A base by suffix interaction was entered as a random slope for the subject random effect but this model did not converge. Base was entered as the random slope because this variable showed more variability by subjects in the group means.

The model showed that the variables base, z=-2.06, p <.05 was a significant predictor of accuracy but suffix, z=.34 was not. For the base variable, when a base word was added to a nonword, the average score decreased .69 points. There was no significant interaction between the base and suffix variable as a predictor of nonword accuracy, z=-1.90. This shows that an English base word in a nonword led children to believe the word was English word. Since there was no effect for suffix this suggests developing readers rely on the base form of the word to make comprehension decisions but not the suffix.

Nonword Response Time

Like the accuracy dependent variable, subjects and items random effects ANOVAs were computed followed by minF′ analyses. A linear mixed model was then computed in R. The mixed model matches and expands on the results of the ANOVAs so only the mixed model results are given in order to not be repetitive.

A linear mixed effects model with a base and suffix interaction was entered as a fixed effect and subjects and word items were entered as random effects. A base by suffix interaction was entered as a random slope for the subject random effect but this model did not converge. Base was entered as the random slope to follow the same parameters as the accuracy model. The results showed that the variable base, t=-2.86, p <.01 was a significant predictor of the speed at which a participant correctly responded to a nonword target but suffix, t=-1.06 was not. For the base variable, when a base word was added to a nonword, the average response time increased 85.38 milliseconds.

There was no significant interaction between the base and suffix variable, t = -.78. These results support the argument that children use the base structure of the word to make comprehension decisions even when that decision eventually led them to reject the word as an

English word. For this task, processing the base form of the word was an inefficient strategy for the children since it led them to delay correctly rejecting the word.

Masked Priming Task

Data from 29 participants was analyzed for English word trials only following the previous two experiments. Accuracy and response times were analyzed with a 2 (Relatedness) X 3 (Condition) factorial ANOVA. Both variables are within factors for the subject analysis and relatedness is a within factor and condition is a between factor for the item analyses. Subject (F1) and Item (F2) variables were computed as random effects variables and minF′ was then computed.

Incorrect responses were excluded from the response time analysis. Response times faster than 150 milliseconds and slower than 2.5 standard deviations above the mean were also excluded. This removed 10.58% of the responses. Table 12 shows the average accuracy and response times for each prime and relatedness category.

For mixed effect model analyses, a maximal model approach was applied for both the accuracy and response time models. Relatedness and condition variables were entered as fixed effects. Subject and word item variables were entered as random effects and given random intercepts. Random slopes were selected first based on the study design and then removed in order to produce a model that converged. Test distributions used for reporting are identical to the previous sections.

Accuracy

Subjects and items random effects ANOVAs were computed followed by minF′ analyses. A mixed model logistic regression was then computed in R. The mixed model matches the results of the ANOVAs so only the mixed model results are given below.

Table 12. Elementary age native English speaker means and standard deviation values for accuracy and response times on the masked priming task.

	Form (castle/cast)	Opaque (corner/corn)	Transparent (hunter/hunt)
Related Prime Accuracy (%)	88.97 (12.05)	88.28 (7.59)	89.66 (12.10)
Related Prime RT (ms)	1404.78 (118.04)	1388.49 (133.94)	1369.21 (98.93)
Unrelated Prime Accuracy (%)	87.93 (11.11)	87.24 (10.98)	89.66 (10.52)
Unrelated Prime RT (ms)	1436.34 (132.75)	1403.97 (109.38)	1412.70 (116.23)

A mixed effect logistic regression was conducted using the binomial link function. A relatedness and condition interaction was entered as a fixed effect and subjects and word items were entered as random effects and were given random intercepts. A relatedness by condition interaction was entered as a random slope for the subject random effect and relatedness was added as a random slope for the word random effect but this model did not converge. Relatedness was then removed from the subject random slope so that condition was assigned as a random slope for subjects and relatedness for word random effects. This model also did not pass the convergence test. Relatedness was then removed as a random slope from the word random effects term and this model did not converge. I then added relatedness as a random slope for both subject and word random effects and this was the most complex model that would converge. The model results showed a no significant effects of relatedness, *z*= .28 or condition, *z*=.67 and no interaction between the two variables, *z*=-.28. This suggests that the children were not using the orthographic, morphological or semantic relationships between the word pairs in the conditions to make accuracy decisions. It also possible that accuracy was high overall and so it is difficult for the model to identify meaningful variability in the dataset.

Response Time

Subjects and items random effects ANOVAs were computed followed by minF′ analyses. A linear mixed model was then computed in R. The mixed model matches and expands on the results of the ANOVAs so only the mixed model results are given.

A linear mixed effect model was entered with relatedness and condition as fixed effects. Subjects and word items were entered as random effects and were given random intercepts. Relatedness by condition interaction was entered as a random slope for the subject random effect and relatedness was added as a random slope for the word random effect but this model did not converge. Relatedness was then removed from the subject random slope and this model also did not converge. Then relatedness was removed from the word random effect and this model passed the convergence test.

The model showed the variable relatedness, t=3.56, p <.001 was a significant predictor of the speed at which a participant correctly responded to a target word. The word pair condition was also a significant predictor, t=2.46, p <.05. There was no interaction between the two variables, t = .59.

A Tukey HSD test was conducted on the factor condition to understand which condition groups were significantly different from one another. The F test was not significant (F (2) =2.06, p=.13) and no conditions were significantly different from one another, but the closest significant difference was between the transparent and form condition, p=.15. The effect in the mixed model and the lack of effect in the F-test is most likely explained by the increased statistical power in the mixed model parameters able to detect differences that the F-test cannot.

Since I am interested in priming effects, I conducted a simple LME test on each condition level with relatedness entered as a fixed factor and relatedness entered as a random slope for both subject and word random effects. These three models did not converge so I removed relatedness as a random slope for the word random effect and this model did converge.

The simple models showed a significant priming effect for the form condition, t=1.96, p = .05, at 35.79 positive priming. There was no significant priming for the opaque condition, t=1.12 at 14.78 ms positive priming and a significant priming effect for the transparent condition, t=2.84, p = .01 at 42.40 ms positive priming. This positive priming pattern matches the adult non-native English speaker group pattern, although there was no significant form priming in the L2 group, and suggests native English-speaking developing readers are still learning to efficiently identify relationships between words during reading.

Vocabulary Group Effects

Vocabulary test percentile score, age and grade level were correlated with lexical decision and masked priming accuracy and response times using a Pearson's correlation test. I also added to two variables: parent reported child vocabulary and reading level (above, at or below their class peers) to the correlation analysis. Age, vocabulary percentile and parent reported child reading level correlated with some experimental dependent variables. Grade and parent reported child vocabulary level did not correlate with any of the dependent measures in the experiment.

Age negatively correlates with the W+S+, the most English-like condition, for English response times. This means the older the participant, the faster they correctly accessed words in that condition. This is most likely explained by the additional reading experience they have gathered through education which has made them better able to easily identify English words that contain two productive morphemes. This skill is not strong enough to give them an advantage in any of the conditions that only has one English morpheme.

Vocabulary percentile negatively correlates with English word conditions with no base word and all of the nonword conditions. This correlation supports the idea that children are using whole word access strategies for English words because there is an advantage only in the conditions where base word processing is not available. In this case high vocabulary readers are quicker to correctly identify no base English words as they are relying on their broader lexical knowledge to make that decision. These correlations are shown in Table 13.

Table 13. Pearson correlation of lexical decision response time and child age, vocabulary percentile and parent reported child reading level.

Lexical Decision Response Times	Age	Vocabulary Percentile	Reading Level
W-S- English	$r = -.267$, $p = 0.16$	$r = -.396$*, $p < .05$	$r = -.527$**, $p < .01$
W-S+ English	$r = -.058$, $p = .77$	$r = -.436$*, $p < .05$	$r = -.491$**, $p < .01$
W+S- English	$r = -.314$, $p =.097$	$r = -.345$, $p = .072$	$r = -.433$*, $p < .05$
W+S+ English	$r = -.373$*, $p < .05$	$r = -.254$, $p = .192$	$r = -.629$***, $p < .001$
W-S- Nonword	$r = -.250$, $p = .19$	$r = -.431$*, $p < .05$	$r = -.642$***, $p < .001$
W-S+ Nonword	$r = -.236$, $p =.22$	$r = -.488$*, $p < .05$	$r = -.595$***, $p < .001$
W+S- Nonword	$r = -.340$, $p = .07$	$r = -.449$*, $p < .05$	$r = -.560$**, $p < .01$
W+S+ Nonword	$r = -.024$, $p = .90$	$r = -.556$**, $p < .01$	$r = -.564$***, $p < .001$

Parent reported participant reading level negatively correlates with response times on all experimental tasks. The higher the parent reported the child's reading level was, the faster they responded to all experimental targets on both the lexical decision task and the masked priming task. This suggests that reading level is a better individual difference measure for these tasks than vocabulary skill. This also supports the idea that pattern recognition is an important part of morphology processing as this is a cognitive activity used during reading that is not related to vocabulary.

In this group, showing skill in lexical retrieval is related to the amount of practice the child has at lexical retrieval and not necessarily the amount of knowledge they have about lexical items. Age and vocabulary percentile did not correlate with masked priming response times. This suggests developing readers are using orthographic and morphological knowledge to make masked priming lexical decisions which are two skills developed during reading that are not derived from vocabulary knowledge. The correlations are reported in Table 14.

Table 14. Pearson correlation of masked priming response time and parent reported child reading level.

Masked Priming Response Times	Reading Level
Unrelated Form Prime	$r = -.367^{*}$, $p < .05$
Unrelated Opaque Prime	$r = -.546^{**}$, $p < .01$
Unrelated Transparent Prime	$r = -.500^{**}$, $p < .01$
Related Form Prime	$r = -.555^{**}$, $p < .01$
Related Opaque Prime	$r = -543^{*}$, $p < .05$
Related Transparent Prime	$r = -.383^{*}$, $p < .05$

Discussion

Elementary age children who are native speakers of English completed a receptive vocabulary test, a lexical decision task and a masked priming task. My hypotheses before the experiment were, first, that elementary age children would be sensitive to word structure changes in the lexical decision task as reflected in their accuracy and response times. This would show they are mentally decomposing complex words during word processing. Second, my goal for the masked priming task was to produce a clear priming pattern in the data since masked priming tasks have not always worked well in previous research. My initial thoughts were that children, like the non-native adult group would show priming in all the conditions. I also hypothesized grade level would influence task performance but because of my small sample size I was only able to test this in relationship in the correlation analyses.

My hypotheses for the lexical decision task were supported and non-native speakers did perform differently when morpheme changes were made to the target word but only for nonword targets. There were no effects for English words. My hypotheses for the masked priming task were supported as each of the conditions showed positive priming and also a response time priming pattern which matched the non-native adult speakers. Additionally, age and vocabulary percentile negatively correlated some lexical decision response times and parent reported child reading level negatively correlated with response times in all conditions for both tasks.

Lexical Decision Task

One other study has used this task with 4th grade children in the United Kingdom and, like me, found no significant effects in English target accuracy although they reported an interaction that was not significant by items (Casalis et al., 2015a). The interaction was identical to the adult native speaker interaction in experiment #1 with a suffix benefit when no base was present but no accuracy benefit to a suffix being added to a base present word.

For my group, it is possible that no effects are present because either accuracy was so high and therefore effects are not detectable or because my sample is not large enough. It is also possible children are using a whole word access strategy in all conditions producing an equivalent response time across conditions. This is likely as children are still learning how to take advantage of morphological decomposition skills during reading. Since I was able to find effects in the native adult group but not in the non-native group which could mean two things. My sample size is small and effects should be clear in any native English speaking group or my sample size is adequate and it is difficult to identify effects in learner groups as Casalis et al. (2015a) found.

For nonwords, the base variable produced lower accuracy and slower response times overall when a base was present in the nonword and the suffix variable produced slower response times when a suffix was present in the nonword. This suggests morphological decomposition is important to unfamiliar or new word processing as the reader does not yet have a lexical entry stored. Again, the base word is the strongest cue, as indicated in both dependent measures, for whether a nonword appears English like or not. Casalis et al. (2015a) found identical main effects for accuracy plus a base by suffix interaction that matches my native adult group where accuracy significantly decreases from the W+S- nonword condition to the W+S+.

Comparing the adult nonword results and the results from Casalis et al. (2015a) to the children group, it is clear this group of readers has not yet learned to take advantage of the suffix structure as a cue (or barrier) to nonword processing. Since there is an interaction in both my native adult experiment and the Casalis child study, it seems like this is a strategy I should find in the response time results for the children group. However, either due to the range of ages and grades in my sample or because my sample is small, there is no evidence child readers are using the suffix structure meaningfully during English or nonword target processing.

Masked Priming Task

There were no accuracy effects for this task possibly because accuracy was high in this group and also the sample size was small.

The response time results of the masked priming task supported my hypothesis for child native speakers of English. The response time results show that the priming task worked in this age group and the sample produced interpretable priming patterns. The priming pattern across conditions matches the other learner group as the opaque condition had the weakest priming effect but the form and transparent condition showed greater priming effects although form priming was not significant in the non-native adult group.

The priming pattern from my study does not match previous masked priming research on this age group. Beyersmann et al. (2012) found both third and fifth graders produced positive priming effects in the transparent condition but negative priming effects in the opaque condition which means participants were actually slower to respond to word pairs like *corner-corn* compared to a pair like *bumper-corn*. Their third graders showed no priming in the form condition and the fifth graders had negative priming the form condition. They concluded their sample processed target words by semantics only.

One explanation for why negative priming was present in their study that could offer insight into my sample's priming pattern is that the children recognize corner and corn are two different words but since the base word spelling is the same it creates interference during processing. But why is this not happening in the form condition with castle and cast? These two also share orthographic overlap and are two different lexical entries. This study also had a much larger sample size than my children group with 40 third graders and 50 fifth graders. This might have also contributed to the different priming pattern in their study compared to my study.

Since some priming effects in their study were negative and mine were all positive in direction, it likely another factor is influencing the different effects found in the Beyersmann et al. (2012) compared to my group. For example, in their transparent word list, they chose to use both inflectional (-ed, -ing) and derivational (-er, -ly) suffixes instead of only derivational suffixes which my study used. Verb endings like -ing could change the priming effect pattern because verbs serve a different grammatical function than nouns and therefore children reading verbs might use a different decomposition strategy than when reading complex nouns.

Castles, Davis, Cavalot and Forster (2007) found that when children first use word recognition strategies, they look for broad relationships and slowly tune these skills to look for more fine-grained relationships when comparing words during reading. This might explain why the children are showing positive priming in the form (castle-cast) condition. Even though it is not a derived word pair, the children in my sample are applying a broad comparison strategy that will most likely become fine grained as they continue their education.

There are two explanations for why children show significant positive priming in the form condition and non-native speaking adults show non-significant positive priming in the form condition. First, they are matching the shared orthography between the words and using their knowledge of spelling to facilitate word processing (Heyer & Clahsen, 2015). A second explanation is that children are removing the base form of the word castle as a decomposition strategy, believing that the words castle and cast are related.

Beyersmann & Grainger (2018) found that when comparing form condition word pairs that had highly productive base words versus non-productive base words, there was positive priming in the productive condition but not the non-productive condition. This suggests form priming is a result of pattern recognition and not a strategy that consists of matching shared orthography or else they would have seen priming in both productive and non-productive conditions. Theoretical research on the process of reading development supports this explanation.

Tamminen, David & Rastle (2015) found that children learn reading skills my generalizing knowledge taken from specific lexical examples. In other words, when a child is presented with a word pair like castle and cast, they are using strategies from previous lexical access opportunities during reading and not just from their knowledge of English spelling.

One reason why parent reported reading proficiency correlated so strongly with task performance is that research has shown morphemic recognition is a skill that is built during reading (Carlisle, 2000). Using morphemes to make sense of words is necessary during independent reading. Previous research has found that grade is the best predictor for morphological decomposition skill (Mahony, Singson & Mann, 2000) but in my sample, the participants may have more educational advantages than other samples of developing readers. Many of my participants reported avidly reading during their free time suggesting that, in my group, reading experience is more of an individual difference that is independent of formal instruction.

Chapter 6. A Quantitative Group Comparison

After giving two identical experimental tasks to three different English-speaking populations, I am interested in whether there are any significant differences in accuracy and response time when all three groups are compared quantitatively in the same model. Below I compare task performance between all three experimental groups: Native English-speaking adults, Non-native English-speaking adults and Native English-speaking children. For both accuracy and response times, I apply a mixed effects model identical to the models applied to the individual group results sections. All models utilize the maximal model approach (Barr et al., 2013) with random slopes selected for random effects that received within factors in the experimental design. Random slope selections are described briefly before reporting the model results to provide information about necessary adjustments made to account for model fit. I included figures that gave new information that was different that patterns identified in the individual experiments. I did not include figures for interactions that were discussed previously. In the general discussion, I will qualitatively compare the results from the individual experiments and from the group comparison model to previous research and interpret what these results mean in relation to word processing patterns. I am able to make these direct group comparisons because each group completed the exact same lexical decision and masked priming task and all word items included in the analysis for each group are identical.

Lexical Decision Task

English Words

For English accuracy, base, suffix and group were entered as fixed effects and subjects and word items were entered as random effects and were given random intercepts. The random effects were separated into two models and effects were compared across models because this was the only model that would converge. These model parameters inflate Type 1 error but allow me to adequately rely on the results. Table 15 shows the average accuracy score in each condition for each participant group by subjects only. The two models shared several significant effects and interactions. Base ($z_{subject}$=2.69, p <.01, z_{word}=2.41, p <.05), suffix ($z_{subject}$=2.95, p <.01, z_{word}=2.69, p <.01) and group ($z_{subject}$=-3.69, p <.001, z_{word}=-3.05, p <.01) were all significant predictors of English target accuracy. When a base word was present, participants were 1.41 points more accurate than when there was no base word in the word and when a suffix was present participants were 1.61 points more accurate than when there was no suffix.

Table 15. Means and standard deviation values for English word accuracy by participant group.

	W+S+	W+S-	W-S+	W-S-	Findings
Native English speaking Adults	97.27 (4.12)	98.66 (2.53)	98.86 (2.71)	97.00 (3.80)	-More accurate than L2 adults and children. -Significant base and suffix interaction.
Non-native English speaking Adults	95.20 (5.59)	97.53 (4.31)	96.76 (5.10)	93.98 (7.29)	-More accurate than children. -No interaction.
Native English speaking Children	91.89 (11.05)	93.65 (7.75)	92.09 (10.22)	91.38 (8.85)	-No interaction.

A Tukey HSD test was conducted to see which of the three participant groups were more accurate than one another. The F value was significant, F (2) = 60.28, p<.001 and showed that all groups were significantly different than one another. The native adult speakers of English were significantly more accurate than the non-native adults and the elementary age children. The non-native adults were also significantly more accurate than the children. This effect is illustrated in Figure 11.

There was a significant interaction between the base and suffix variables, $z_{subject}$=-3.65, *p* <.001, z_{word}=-3.04, *p* <.01 the suffix and group variable $z_{subject}$=-2.10, *p* <.05, z_{word}=-1.99, *p* <.05 and the base, suffix and group variables, $z_{subject}$=2.37, *p* <.05, z_{word}=2.10, *p* <.05.

For the base and suffix interaction, a simple model was conducted on each level of the base variable with suffix as both a fixed factor and a random slope for the subject random effect. These two models did converge.

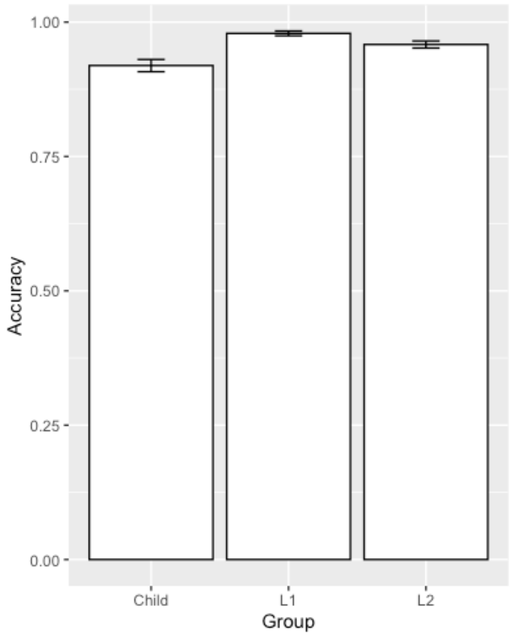


Figure 11. English target accuracy on the lexical decision task by participant group.

The pattern showed that when a base word was present all participants were not more or less accurate when the suffix was and was not present in the word, *z*=-1.42. When a base word was absent, all participants were significantly more accurate when a suffix was in the word, a .80 point advantage, compared to when there was no suffix, *z*=1.99, *p* <.05.

An important note for the group and suffix interaction is that the means among the three groups for the suffix factor do not appear to portray an interaction as depicted in Figure 11. I changed the parameters of the fixed and random effect model in different ways to see if the significant interaction was a result of an error. The interaction was not significant in the result summary when the variable base was removed as a fixed factor in the model and it was also not significant in an ANOVA test. However, in the mixed model test when base was entered as a fixed factor, a group and suffix interaction was always significant regardless of the random effects structure entered (random intercepts only, random slope and random intercept, random slope only and no random intercept). Since there is no significant interaction in the ANOVA test and no significant interaction in every version of the mixed model without the base variable as a fixed effect, this interaction may not be reliable.

For the group and suffix interaction, a simple model was conducted for each group with suffix as both a fixed factor and a random slope for the subject random effect. These three models did converge. The models showed no main effect of suffix for any of the three participant groups, but the slope did vary in the learner groups compared to the native English group. When a suffix was present in a word, native English speakers earned .26 more points compared to a suffix absent word. For the non-native adults and elementary children, when a suffix was present in an English word, both groups were less accurate earning .02 and .01 fewer points respectively. This interaction is depicted in Figure 12.

For the three-way interaction, a simple model was conducted for each group with a base by suffix interaction as a fixed factor and base as a random slope for the subject random effect. These three models were the most complex that would converge. The models showed a significant base and suffix interaction only in the native English-speaking group, z=-2.49, p <.05 but not the learner groups.

This is the same pattern identified in the overall base and suffix interaction where adding a suffix to a no base word improves accuracy significantly but there is no accuracy advantage for a suffix present in the base present condition (W+S+ bony).

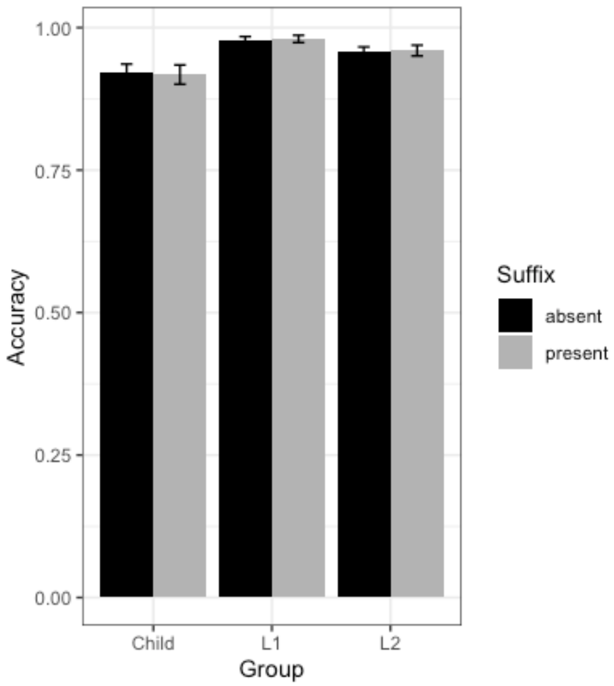


Figure 12. English target accuracy for the suffix variable by participant group.

Next, a linear mixed effects model was conducted for the English word response time values. Table 16 depictes the average response times for each condition by participant group by subjects only. Base, suffix and group were entered as fixed effects and subjects and word items were entered as random effects and were given random intercepts. The base and suffix interaction was entered as the random slope and this model did converge. The model showed that the variables base, *z*=.92 and suffix, *z*=-1.15 were not significant predictors of English target response times but group was, *z*=-9.30, *p* <.001. There were no significant interactions.

A Tukey HSD test was conducted to see which of the three participant groups were faster to respond to English words than one another. The F value was significant, F (2) = 1076, p<.001 and showed that all groups were significantly different than one another. Native English-speaking adults were significantly faster to respond to English targets compared to the non-native and children group. The non-native group were also significantly faster to respond to targets than the elementary group. This effect is depicted in Figure 13.

Table 16. Means and standard deviation values for English word response times by participant group.

	W+S+	W+S-	W-S+	W-S-	Findings
Native English speaking Adults	697.96 (81.70)	671.86 (81.76)	682.79 (79.05)	685.74 (77.92)	-Faster than L2 adults and children. -No interaction.
Non-native English speaking Adults	738.51 (129.25)	711.02 (129.93)	715.63 (121.66)	725.77 (110.28)	-Faster than children. -No interaction.
Native English speaking Children	967.95 (149.52)	961.27 (136.13)	978.48 (150.98)	980.20 (149.47)	-No interaction.

These results indicate that the interaction between the base and suffix variables is mainly derived from the adult native group experiment. The suffix by group interaction also supports this as it shows the slope changes between the native adult group and the two learner groups. When a suffix is present in an English word, the native adult speakers benefitted from the structure whereas the learner groups showed a decrease in accuracy. This supports the argument that non-native adult speakers of English and developing English readers do not use suffix decomposition as a strategy for word comprehension. The lack of response time effects in the English target words indicates that regardless of what strategy the participant using to make a lexical decision, that strategy does not facilitate or inhibit the decision-making process.

Nonwords

Base, suffix and group were entered as fixed effects and subjects and word items were entered as random effects and were given random intercepts for the accuracy model. Suffix was entered as the random slope. Table 17 displays the average nonword accuracy for each participant group by subjects only. The model showed that the variable group, z=-2.37, $p < .05$ was significant predictor of nonword accuracy but base and suffix were not.

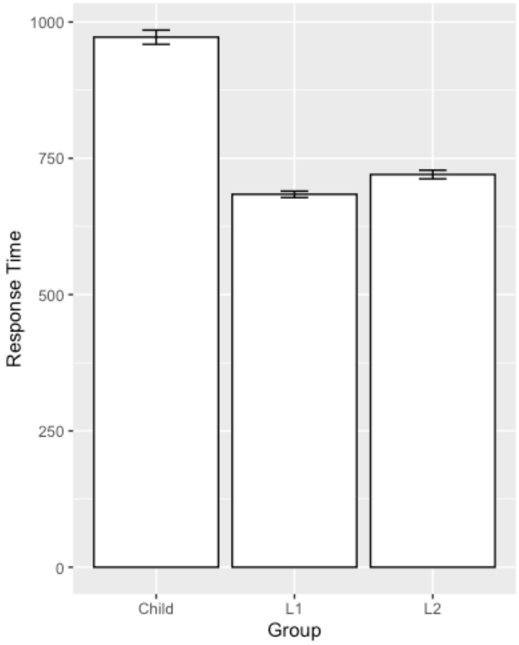


Figure 13. English target response times by participant group.

There was also a base and suffix interaction, z=-3.03, $p < .01$ but there were no other interactions. The group pairwise comparison showed that native speakers were both significantly more accurate on nonword targets compared to the non-native adults and children but there was no difference between the L2 adults and children, F (2) = 82.04, $p <.001$. This is depicted in Figure 14.

For the interaction, when there was no base word, adding a suffix did not affect accuracy but when there was a base word, adding a suffix significantly decreased word accuracy, z=-4.72, $p < .001$. This matches the interactions reported in experiment #1 and #2.

For nonword response times, base, suffix and group were entered as fixed effects and subjects and word items were entered as random effects and were given random intercepts. A base and suffix interaction was entered as the random slope for subjects and converged appropriately. Table 18 shows the average nonword response times for each participant group by subjects only.

Table 17. Means and standard deviation values for nonword accuracy by participant group.

	W+S+	W+S-	W-S+	W-S-	Findings
Native English speaking Adults	74.73 (15.14)	91.36 (8.95)	93.55 (7.97)	92.27 (7.06)	-More accurate than L2 adults and children. -Worst accuracy in W+S+ condition.
Non-native English speaking Adults	67.47 (25.33)	78.27 (21.59)	83.98 (19.97)	82.45 (17.44)	-As accurate as children. -Worst accuracy in W+S+ condition.
Native English speaking Children	66.94 (18.43)	78.45 (15.07)	88.12 (15.87)	83.97 (14.90)	-No interaction.

The model showed that group, *z*=-6.02, *p* <.001 was a significant predictor of response speed but no other factors were significant. All groups were significantly different than one another for nonword response time speed. The native adult group was significantly faster than both learner groups and the non-native group was significantly faster than the children group. This effect is depicted in Figure 15.

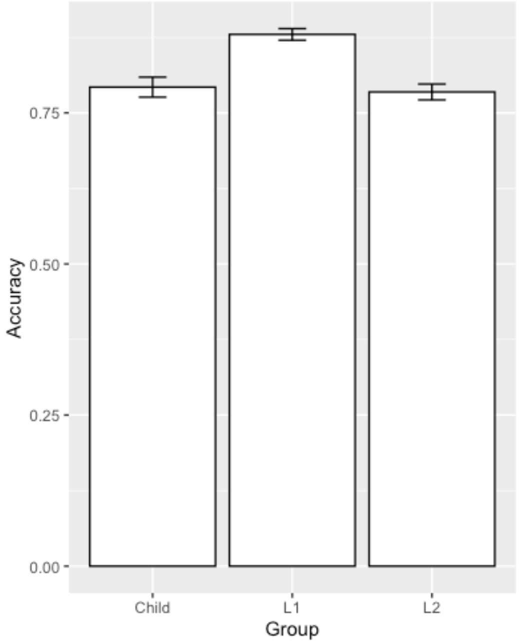


Figure 14. Nonword target accuracy by participant group.

There was also a significant interaction between the base and suffix variable as a predictor of nonword response times, $z=-2.72$, $p<.01$. The simple models showed that when there was a base word adding a suffix to the nonword significant slowed the response time for all participants, $z=-3.98$, $p<.001$ but there was no response time change in the suffix condition when there was no base word in the nonword.

Table 18. Means and standard deviation values for nonword response times by participant group.

	W+S+	W+S-	W-S+	W-S-	Findings
Native English speaking Adults	901.94 (123.90)	830.18 (131.73)	802.28 (116.98)	799.21 (118.15)	-Faster than L2 adults and children. -Slowest in W+S+ condition.
Non-native English speaking Adults	961.61 (228.59)	956.53 (245.73)	907.81 (193.92)	885.03 (197.38)	-Faster than children. -No interaction.
Native English speaking Children	1240.50 (233.75)	1167.00 (183.97)	1111.42 (197.11)	1083.55 (175.75)	-No interaction.

The nonword group comparison shows that all groups had the most difficulty with the W+S+ nonword condition which has both an English base word and English suffix. This is supported by the interactions in both the accuracy and response time results. For response times there was also a main effect of base word which supports the argument that the base word structure is the morpheme that provides the best cue for lexical decisions. Finally, the group main effect indicates that both learner groups, adult and child were equally as bad as at making accurate nonword decisions which shows that non-native adult speakers of English, even after years of experience with the language, use lexical decision strategies for unfamiliar words that are more akin to a fourth grade native English speaker than an adult native speaker.

Masked Priming Task

Relatedness, condition and group were entered as fixed effects for the accuracy model. Subjects and word items were entered as random effects and were given random intercepts. Condition was assigned to be the random slope for the subject random effect, and this was the most complex model that would converge. The model showed that group was a significant predictor of target accuracy, $z=-8.93$, $p <.001$ but there were no other significant predictors or interactions. The elementary age group was significantly less accurate compared to both of the adult groups but the two adult groups were not significantly different than one another, $F(2) = 123$, $p < .001$. This is depicted in Figure 16.

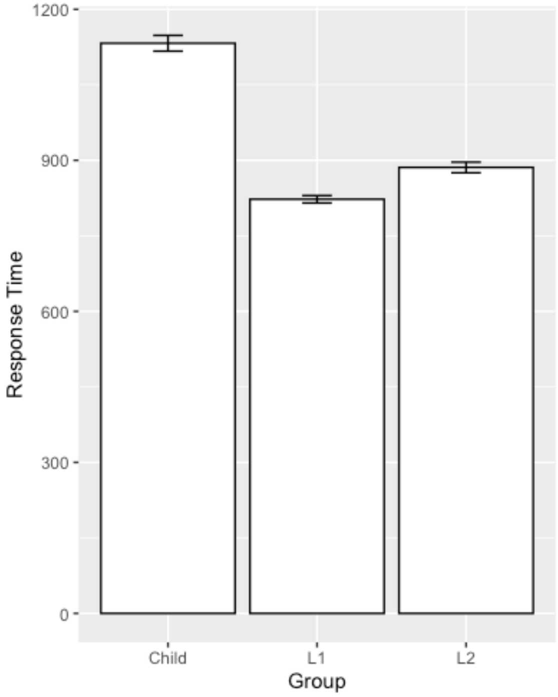


Figure 15. Nonword target response time by participant group.

Relatedness, condition and group were entered as fixed effects for response times. Subjects and word items were entered as random effects and were given random intercepts. Condition was assigned to be the random slope for the subject random effect, and this was the most complex model that would converge. Table 19 shows the average response time priming effects for each participant by subjects only.

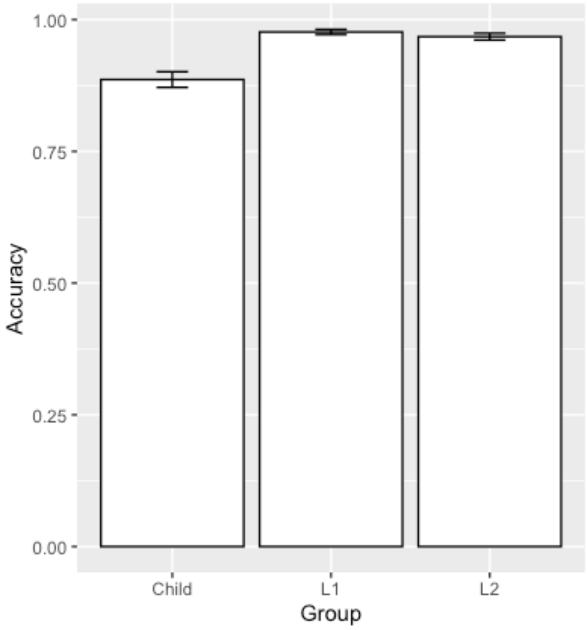


Figure 16. Target accuracy on the masked priming task by participant group.

The model showed the variables relatedness, *t*=-2.24, *p* <.05 and group, *t*=-6.39, *p* <.001 predicted target response times but not condition. When a target word was assigned to a related prime, participants were significantly faster to correctly respond to the target word. All groups were significantly different than one another, *F* (2) =403.30, *p* <.001.

Table 19. Means and standard deviation values for response time priming effects in each condition by participant group.

	Form	Opaque	Transparent	Findings
Native English speaking Adults	3.89 (53.26)	11.52 (64.91)	50.40 (56.39)	-More accurate than children and as accurate as L2 adults. -No other effects.
Non-native English speaking Adults	16.11 (69.36)	2.95 (60.44)	40.11 (58.02)	-More accurate than children. -No other effects.
Native English speaking Children	31.57 (65.79)	15.48 (90.63)	45.37 (61.55)	-No effects.

The native adult group was faster to respond to targets than both learner groups and the non-native adult group was faster than the children. This group main effect is depicted in Figure 17.

There were significant interactions between relatedness and condition, *t*=3.68, *p* <.001 and relatedness, condition and group, *t*=-2.02, *p* <.05. To test the relatedness and condition interaction, simple models were conducted on each condition level with relatedness entered as fixed factor and random slope for both random effects. The models showed significant positive priming in the transparent condition, *t*=6.27, *p* <.001 but in no other condition.

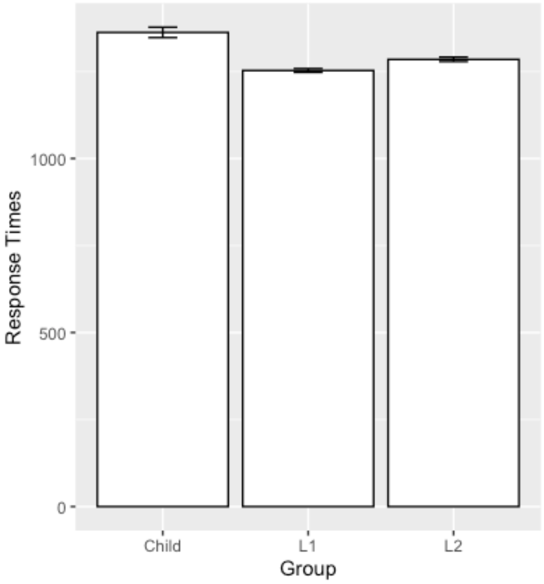


Figure 17. Target response times on the masked priming task by participant group.

For the three-way interaction, a relatedness and condition interaction was entered as a fixed factor for each participant group, with condition as a random slope for subjects. The models showed a significant interaction between relatedness and condition for both adult groups but not the children participant group. These interactions match the data from the individual experiment results section.

These results show that all the participants benefitted from the relatedness condition in the word pairings. This was shown in the main effect of relatedness in the response times. There was a significant interaction between relatedness and condition which was driven by the significant positive priming effect in the transparent condition. This means that all participants were most likely using the same strategy to process the transparent word pairs like hunter and hunt. The relatedness, condition and group interaction showed that this interaction was mainly isolated to the adult experiment groups. This is most likely because my sample size in the third experiment was not large enough.

Chapter 7. General Discussion

Three English-speaking populations, adult native speakers, adult non-native speakers and elementary age native speakers completed the same lexical decision task and masked priming task.

For the lexical decision task, the English word effects were in the adult native speaking group in the accuracy dependent measure only. The non-native adult group featured response time differences between reading age and vocabulary groups, but these were the only response time effects in the learner groups. Accuracy gives information about how English word morphology cues word knowledge but does not give information about whether the morphemes cue more efficient or less efficient visual word processing. Therefore, I can say that the L1 adults decomposed English word morphemes during visual word processing, but I cannot say whether this strategy was used to access the word more quickly.

For the nonword targets on this task, effects were consistent across groups. Both adult groups showed lower accuracy in the W+S+ condition compared to the W+S- condition resulting in a base and suffix interaction. In this case having two English cues is damaging to correct nonword rejection.

There were a few accuracy interactions in the non-native English group when vocabulary group was entered as a fixed factor. English morphemes, base or suffix, when present in nonwords were damaging to accuracy in low vocabulary participants. In this case, the familiarity of the English structures led low vocabulary participants to believe the nonwords were English words. A three-way interaction also showed high vocabulary non-native adults matched the nonword accuracy pattern in the native adult group but the low vocabulary group did not. This provides evidence that readers are able to ignore misleading cues with increase vocabulary experience. Finally, children were significantly less accurate and slower to respond to nonword that featured a base word. This shows they are using the base word structure during new word processing.

For the masked priming task, native adult speakers were the only group that showed a significant interaction between relatedness and condition. In the group comparison in chapter 6, there is a significant interaction for non-native adults as well. This interaction is driven by the transparent condition positive priming in both groups. The children showed main effects of the two variables but no interaction. Again, this is likely due to the sample size as there was significant positive priming in both the form (almost) and transparent condition. The interaction for non-native adults was not present in the results for experiment 2 because a third factor was included in the model for the initial analysis whereas the group comparison chapter only included two factors for each group thus giving the model more statistical power.

The priming effects from the masked priming task provide an interesting perspective to how these groups are processing the two relatedness conditions. Positive or negative priming occurs when the effect of the unrelated condition minus the related condition is significantly different than zero. Positive priming would mean that the reader is processing the related condition faster than the unrelated condition and negative priming would mean that the reader is processing the unrelated condition faster than the related condition. For this research, theories about morphological decomposition could potentially explain why there are priming effects in some conditions but not others.

The native adult group had significant positive priming effects in both the transparent (hunter-hunt) and opaque (corner-corn) condition which matches previous research on this topic (Rastle et al., 2008). This pattern suggests the adult native group engaged in obligatory decomposition during visual word processing because they showed significant priming in both conditions that had English suffixes one of which shared a semantic relationship between the prime and target pair and one that did not. Since the reader still shows a response time benefit from the related prime-target pairing that does not share a semantic relationship, I know they are not using semantic relationship as a source for morphological decomposition.

In the learner groups, the transparent condition showed significant positive priming and significant positive form priming for children and non-significant positive form priming for non-native adults. The fact that there is no significant positive priming in the opaque condition for both the non-native adult group and children group, indicates there is either a whole word strategy being used or interference is occurring during morphological decomposition. It also shows that both groups are using a different strategy to compare form condition word relationships as evidenced by the positive priming in the form conditions. This strategy is either an orthographic comparison strategy due to not yet tuning their pattern recognition to a more efficient method (Castles et al., 2007) or due to base word stripping which is when the reader removes the base word from the word and uses it for comprehension purposes. In this case, to identify a relationship between the prime and target pair.

It is also important to consider how the individual difference measures gathered in each group influence task performance on both the lexical decision and masked priming task. These measures include vocabulary proficiency for all three groups, age of acquisition for the L2 group and grade and parent reported reading level for the elementary age children. Briefly, it was clear that vocabulary proficiency plays some role in accuracy for the English words and response times for the nonwords for the native adult group in the lexical decision task. It did not influence priming effects on the masked priming task.

Age of English reading acquisition was significant on four of the six dependent measures in the non-native adult group and vocabulary proficiency was significant in all of the dependent measures. This means that vocabulary proficiency is a better individual difference measure for non-native adult speakers of English.

Parent reported child reading level negatively correlated with all response time measures on both tasks and vocabulary percentile score negatively correlated with six of the eight lexical decision conditions. This suggests that for children, reading experience is a better individual difference measure than vocabulary skill.

These results partially support my hypotheses for the lexical decision task. Both learner groups did not use the English morpheme structures to guide their English lexical decisions but they did use the structures for nonword decisions. My hypotheses were supported for the masked priming task with the exception of the opaque priming condition predictions in the learner groups and the form condition in the L2 group. Individual difference predictions were partially supported as well. Every experiment showed some relationship with an individual difference measure either in the correlation analyses or as a factor in the models. However, individual difference measures did not seem to influence masked priming behavior which has been reported in other studies (Andrew & Lo, 2013; Heyer & Clahsen, 2015).

Qualitative Group Comparison

Adult Groups (L1 and L2 Group)

Lexical Decision Task

The quantitative group comparison revealed that for English word accuracy, the native adults benefitted from having a suffix present in the word but non-native adults did not. In the individual experiments, the two adult groups showed different effects on the English target trials with native adults having three main effects and two interaction for accuracy and no response time effects. The non-native adults had no accuracy or response time effect besides a performance advantage in both dependent measures for more proficient English speakers. This means I cannot really compare the two groups based on their performance on this task except to say that it is clear the native speakers are processing the morphological structure of the English words and using strategies that are helping them make lexical decisions. Specifically, base word processing and in the low vocabulary group and the entire sample used a suffix cue when no base word was present. It is not clear the non-native group is using the morphological structure of the words to make lexical decisions and more research is needed to identify when and how non-native English speakers use English morphemes during processing. There were also no response time effects for English word items in Casalis et al. (2015b), another study that tested non-native English speakers with this task and materials that had a much larger sample size than my non-native group.

Vocabulary proficiency does affect how a native reader processes morphemes, as low vocabulary readers benefit especially from a suffix cue when no base word is present in the word. This means that having no base word to process is especially harmful to low vocabulary reader accuracy.

The nonword target results give a more defined example of how readers process new, unfamiliar complex words. There were no group interactions in the nonword analyses but the non-native group were slower to respond to nonword targets than native adults. But the main effects and individual group experiments show both the native speaker and the non-native English speaker are using English structures in nonword to make recognition decisions. The adults are not processing nonwords based on their whole word representation but based on the components that make up the word. In this case, having two English cues in the same nonword leads to the wrong decision or incorrectly identifying the word as an English word. This was shown by having the worst accuracy and slowest response times in the W+S+ condition in native speakers and the worst accuracy in this condition in non-native speakers. We do not see response time effects in the L2 group mainly because of the low accuracy rate in this group causing many trials to be excluded for the response time analysis.

The effects in both the L1 and L2 adult group match the nonword results of the French-English bilinguals in Casalis et al. (2015b). In their study, there was a main effect of base and suffix with base present and suffix present words producing poorer performance than nonwords without these structures and also an interaction in that the W+S+ condition produced the poorest performance. The results from my study and the results from Casalis et al. (2015b) combined suggest that using morphological structures to make decisions about knowledge of unfamiliar words is a common skill among all English readers with English reading experience that is equivalent to my participant groups. My L2 group had an average age of acquisition at 6.5 years of age.

Additionally, there were three interactions in the nonword accuracy model for non-native speakers using vocabulary proficiency as a fixed factor. These interactions showed low vocabulary readers performed worse whenever there was an English structure in the nonword, base or suffix and also high vocabulary readers process English morphemes similarly to native speakers with the worst accuracy in the W+S+ condition. This means that while low vocabulary non-native readers are tricked by any kind of English structure, the high vocabulary non-native readers are only showing clear accuracy decreases when two English structures are present in the nonword. In my study, vocabulary proficiency influences native adult recognition of English targets and non-native recognition of nonword targets.

Masked Priming

Accuracy was high overall in both the L1 and L2 groups and there were no accuracy effects in the L1 adult group and only a main effect of reading age and vocabulary group in the L2 adults.

Rastle et al. (2004) mentions there were no accuracy effects in their masked priming study of native English speakers and Li et al. (2017) found no accuracy effects in their monolingual comparison group. Li et al. (2017) did find accuracy effects in the priming effects of two Chinese-English bilingual groups separated by low and high English proficiency. The high proficiency bilinguals in their study showed consistent positive response time priming in all three conditions, transparent, opaque and form and the low proficiency bilinguals showed consistent positive priming in the transparent and form condition only.

In my study, for response times, the L1 group had an interaction between the relatedness and condition variables which was driven by the transparent priming effect. In the L2 adult group, there was a main effect of relatedness and condition but no interaction in the individual experiment but an interaction in the group comparison simple effects analysis. There was also a significant priming effect in the transparent condition. Reading age and vocabulary group had no influence on masked priming effects beyond an overall speed advantage in the more proficient groups. To make sure individual difference measures did not interact with priming effects, I entered vocabulary group and reading age and vocabulary group into the respective L1 and L2 response time priming effect models. No individual difference measure for either group interacted with relatedness for any masked priming condition. These results do not match

Andrews and Lo (2013) who found that high vocabulary participants showed more priming on the transparent condition than the opaque condition.

For response time priming patterns, the L1 adult showed a declining positive priming pattern that matched the Rastle et al. (2008) priming patterns with the transparent condition producing the largest positive priming effect (48 ms), followed by the opaque condition (10 ms) and then the form condition which showed positive priming that was not significantly different than zero (3 ms).

The L2 group had a transparent priming effect of 40 ms followed by a form priming effect of 16 ms and then an opaque priming effect of 3 ms. If the L2 readers were to respond like native speakers, the form condition should show the least amount of priming not the opaque condition. The priming pattern across conditions is different between adult groups. Li et al. (2017) also found a similar priming pattern in the low proficiency L2 group when compared to my L2 group which shows that this priming pattern may have some meaning in relation to how non-native adults apply lexical access strategies for complex and complex appearing English words.

What Li et al. (2017) hypothesizes is that native English speakers use the English suffix ending as a strategy for efficient processing and "turns off" the part that of the brain that recognizes the two words are not semantically related. The L2 adults are aware that the opaque and transparent condition are different as evidenced by their response times, but they are not using the suffix as a tool for quick lexical access like the native speakers. It is clear some type of interference is taking place, but it is unclear from these results what is the source of that interference. Future research should address what causes non-native adult speakers to abandon a morphological decomposition strategy in the opaque condition. This lack of effect could be an artifact of the words in the condition but since both of my statistical models accounted for this variability, this explanation is less probable.

The opaque condition has caused difficulty in other L1 adult research studies. In Diependaele et al. (2005), participants showed significant positive priming for both visually and auditorily presented primes in the 67 ms presentation priming duration but a negative priming effect in both modalities the 40 ms priming duration. This means that visual word processing strategies vary depending on how much time the reader has to make sense of the word. Finding a relationship between corner and corn may take more time for it be a facilitative process which may be why the authors found the effect only in the 67 ms window. If it does take more time to apply a morphological decomposition strategy in this condition, this may explain why I did not find an effect in my non-native English-speaking adults. Perhaps a longer prime duration would reveal significant priming effects in this group in this condition.

The form condition produces non-significant positive priming in my L2 group and significant positive priming in the low proficiency group in Li et al. (2017). This suggests, with a larger a sample size, L2 participants similar to mine may see a relationship between castle-cast and are using that relationship to make a lexical decision. The native speakers do not see the two words as related which why we do not see any facilitative effect for them, significant or non-significant, in the form condition. This relationship is either driven by orthography at the pre lexical level (Heyer & Clahsen, 2015) or by base word stripping at the lexical level of processing (Grainer & Beyersmann, 2017).

Vocabulary Learners (L2 and Child Group)

Lexical Decision Task

I cannot compare the two learner groups on English target words in the lexical decision task qualitatively because there were no accuracy or response time effects in the child group. In the quantitative group comparison model, the slopes for the suffix variable show that both learner groups were less accurate when a suffix was present in the word. This was indicated by the suffix and group interaction. This suggests that both learner groups are not confident with how to use suffix information to make lexical decisions. This also could indicate that the readers are unclear of the suffix meaning in terms of its function for word formation.

There are a couple explanations for why there are no effects in the child group. One is that accuracy is so high that there are no detectable differences based on condition but since both the adult groups showed effects on English target words and they had high accuracy rates, it is more probable that the lack of effects was due to a small sample size. Another reason is that children are not yet practiced in morphological decomposition as a strategy for word recognition and are still learning how to master it. Casalis et al. (2015a) only found a main effect of base word for response times in their study which showed that 4th graders were significantly slower to respond to words with a base than those without one. Their study had 40 participants which shows that a larger sample size still might not provide an opportunity to detect more effects.

Finally, is possible that each reader in the child group uses morphological decomposition differently and any patterns among the group are undetectable. For example, some developing readers might use base word recognition to identify an English word and some might use suffix recognition as a cue for English word recognition. If readers in this age range are using morphological decomposition strategies during word reading but they are all using different ones, then those individual, reader-based patterns would wash out in the analysis. Since, I accounted for subject variability in the statistical models this also might not be the best explanation.

For nonwords, the combined group model showed that the two learner groups had the same accuracy for nonword targets. There were no interactions with group. In the individual experiments, the non-native adult group used the base word forms and the combined base and suffix morphology to make nonword recognition decisions, but these strategies did not affect their response times. This is also likely because of the excluded trials in the RT dataset. The child group used the base word to make nonword recognition decisions and during correct nonword rejections as evidenced in the response times. These comparisons relate to how the groups differ in terms of how English morphology makes their decision more or less efficient. The L2 group used both structure types to help decide if the word was an English word or but it did not affect response speeds. The child group used base word morphology to help decide if the word was an English word or not and this made their correct rejections less efficient.

Neither group is using the suffix to help or hurt them in their nonword decision making but the base present nonwords do cause decreased performance for both groups. Casalis et al. (2015b) also found that bilinguals slowed down their correct rejections to nonwords when a base was present but not for a suffix present nonword. The Casalis et al. (2015a) study with 4th graders found no effects for nonwords.

It appears that morphological decomposition is a skill children use consistently when they encounter a nonword (or new word) but not an English word. Since the sample size is the same for the English and nonword analyses, clearly children are using a morphological decomposition strategy differently when reading a familiar word versus an unfamiliar word. As the English word results do not match the null nonword results from the Casalis et al. (2015a) study, the nonword analyses was able to add to this body of research.

Masked Priming Task

Neither of the learner groups show accuracy effects for this task except for the L2 individual difference groups main effects. Both learner groups were faster to respond to target words that were paired with a related prime which means that both groups were sensitive to the relatedness factor during word recognition. For the child group this is a good indicator the participants are experiencing the task in a similar manner to the other groups.

Both learner groups showed a main effect of condition, but the Tukey HSD F-test was not significant for either group. This means that the response time differences between the overall task conditions were not different enough to produce significant results. It is also possible the mixed model is able to detect differences that the F-test is not able to. The quantitative group comparison revealed a significant interaction between relatedness and condition for the adult learner group but not the child group. This may be because my sample size was larger in the adult group, but it also could be a result of the adults having a longer duration of English exposure compared to the child group. The average age in the child group is 10.2 years and the average years of English exposure in the adult group is 16 years.

The dependent variable of interest for this task is the response time priming effects. Significant positive priming occurred in both the form and transparent conditions for children and significant transparent priming occurred in the L2 adult group. Non-significant positive priming values were weaker in the opaque condition in both groups.

Coupled with the same priming pattern in the Li et al. (2017) study, but no the same effects, this similarity across the two groups and across my study and the Li et al. (2017) study shows that learners of English vocabulary process derived words like *corner* more slowly than native speakers and this could be a result of difficulty reconciling the -er ending with a word like corner during morphological decomposition.

Beyersmann et al. (2012) found negative priming effects on a similar masked priming task in 3rd and 5th graders and hypothesized that the lack of priming in the opaque condition is due to the readers still learning how to mentally decompose falsely derived words like *corner*. However, while this may explain processing patterns in new reader groups, like children, it does not explain why L2 adults were unable to use the opaque condition suffix as an efficient tool for quick word processing like the transparent condition.

The participants in my non-native adult group had been reading English for an average 15 number of years. It is unlikely the majority of these participants are still learning how to morphologically decompose falsely derived words and it is more likely that the strategy they elect to use is inefficient. L2 adult speakers may have learned to abandon morphological decomposition for falsely derived words and use a whole word access strategy like they would for a non-suffixed word or morphological decomposition in this condition is chronically slow for these types of words in this reader group.

Beyersmann et al. (2012) did not find positive priming in the form condition in the 3rd and 5th graders and they conclude this is because their readers did not use the shared orthographic structure between words like castle-cast to produce a more efficient word access strategy. However, I did find positive priming in the form condition in my study with elementary age children and non-significant priming in the L2 adult group. Li et al. (2017) found significant positive priming in the form condition for the low proficiency L2 group and was unable to explain why positive priming occurred in the form condition but not the opaque condition for their study. I also interpret this as a confusing pattern. If the learner groups benefit from orthographic overlap between words in the form condition, why is there no benefit in the opaque condition?

Beginning with the L2 adult group, in the form condition, participants see the words *castle* and *cast* as related based on their shared orthography and this produces a faster response time than the pairing *maker* and *cast*. In the opaque condition, participants apply a morphological decomposition strategy to *corner* but then when *corn* presents on the screen, the shared orthography but lack of semantic relatedness creates interference which produces an equivalent response time to the unrelated opaque condition. An alternative explanation for this condition is that the same strategy is being used to process both related and unrelated opaque prime words. In the transparent condition, participants apply a morphological decomposition strategy to *hunter* and when *hunt* appears, the semantic relationship between the words facilitates word access producing faster response times than the unrelated transparent word pairings.

It is possible there are two different word processing strategies occurring in both learner groups in this study because the only shared similarity across conditions is there is matching orthography in the related pairings. However, the reader is not applying an orthographic similarity processing strategy to all conditions because then all priming effects would be equal across conditions. They are also not using a morphological decomposition strategy across all conditions because there is no suffix ending in the form condition and this condition shows positive priming effects in both groups. It is also possible the children are using the same strategy, orthographic matching, across all three conditions and the weaker positive priming number in the opaque condition is due to interference as they recognize there is no semantic function of adding -er to a word like *corner*.

Individual differences did not affect processing strategies in this task for either group. This suggests that either my learner groups were equivalent in vocabulary knowledge within groups or skills that build vocabulary proficiency do not relate to morphological processing on this task.

Native English Groups (L1 and Child Group)

Lexical Decision Task

As with the learner groups comparison, I cannot compare the two native English-speaking groups on the English target words because the child group did not show any accuracy or response time effects. My ideas about why this group did not show any effects is reported in the previous section. Since children are all at different stages of reading experience between grades 3 and 6, it is likely they are all using different word decomposition strategies compared to the native English-speaking adults who have had many years of reading practice to develop consistent complex word processing strategies. Additionally, I tested college undergraduates for my native adult group who most likely are all highly capable and skilled in reading. Children are not only practicing how to read efficiently but the range of variability in reading skills in elementary school is most likely larger than in college age students.

The suffix by group interaction in the quantitative group comparison showed that the native adults were more accurate for suffix present English words and the native children were less accurate for suffix present English words. This shows that children have not yet learned how to use the suffix information to make more accurate English lexical decisions.

There are some similar patterns across groups for the nonword targets. For accuracy and response times, both groups saw a decrease in accuracy when a base word was present. The adults were affected by both base and suffix present nonwords whereas the children were only affected by base present nonwords. This difference is likely caused by the children not using suffix processing when encountering new words.

These results suggest that base word recognition is a cue that aids in word recognition while developing English reading skills and is still used even when those skills are solidified during adulthood. Also, reading nonwords with English base words creates an inefficient word search as both the developing reader and experienced reader decides whether they do in fact know the word as an English word.

Masked Priming

Neither group had any accuracy effects on the masked priming task. For response times, the native adult group had an interaction between the relatedness and condition variable whereas the children only showed main effects in these two conditions. This makes sense from a developmental perspective as it shows adults are using a more targeted decomposition strategy whereas the children are merely identifying general relationships between words. It is also possible there is no interaction in the children group because of the small sample size.

The priming effect pattern between groups shared some similarities and some differences. Both the transparent and opaque priming effects were similar across native English-speaking groups. Adults displayed a 48 ms positive priming effect in the transparent condition while children showed a 42 ms positive priming effect. In the opaque condition, adults showed a positive priming effect of 10 ms and children averaged a 14 ms advantage compared to the opaque unrelated condition. In the form condition there are clear differences between the two groups. The adult group shows a priming effect of 3 ms which is not different than zero and the child group shows a priming effect of 35 milliseconds.

In the previous section, I compare the opaque condition priming effects between the learner groups. I found while the overall pattern among the conditions was similar across learner groups, interference occurring in the opaque condition is weaker in the children group. As I compare the two opaque conditions between the native speaker groups, it is clear the numbers are more similar than the numbers in that condition in the learner group.

Li et al. (2017) suggests that opaque positive priming occurs in their native English - speaking group because they have learned to ignore the falsely derived relationship. With this idea in mind, it is possible that both native English-speaking groups in my study have learned to ignore the false derivation and apply a suffix segmentation strategy for efficient processing. It is also possible that both groups experience interference in this condition as suggested in the previous section. Interference is possible in the adult group because the positive priming effect in this condition is weaker than the transparent condition which is a condition that would utilize a morphological decomposition strategy as well.

One explanation for why the transparent priming effect in both native English-speaking groups in this study is greater than the opaque condition is that the semantic relationship between the prime and target pair creates a boost in processing speed when responding to the target item. Since the opaque condition does not have a semantic relationship between the prime and target pair this boost is absent in the priming effect patterns.

The two alternatives for why weak or non-significant occurs in the opaque condition, interference or whole-word access, suggests two different trajectories during reading development. First, if less priming occurs because of interference during obligatory decomposition, this means that increased practice and English exposure would eventually correct or lessen interference leading to larger positive priming effects in adult reading. Opaque priming effects were significant in the ANOVA test for the native adult group but were not for the children group which supports this theory but this also could have been due to sample size differences as well. The second alternative suggests that readers adopt different word processing strategies at different learning periods and that interference or whole word access may be occurring in the learner groups as a result of their reading experience or non-native status which may or may not correct itself with more reading experience.

Interpreting the Results

The results from the individual studies and the combined group model show that all groups use base word processing to make lexical decisions and the adult groups especially benefit from this strategy. This suggests base word processing is a strategy adopted either earlier in reading development or it has a privileged status in the visual word recognition system. This claim is supported by the lack of suffix effects in the child group and the strong form priming effects in the learner groups. This may be because suffix meaning in relation to word meaning is not always regular and takes more reading experience to learn and use. However, all groups were able to show strong priming in the transparent condition in the masked priming task which means they all understand how derivational suffix formation works and are able to take advantage of this knowledge.

Results from the masked priming task indicate that morphological decomposition strategies are used differently across age and native language status and support for obligatory decomposition is only present in the native speaker groups. Although both learner groups matched the form word pairs to produce positive priming effects, but not significant in the L2 group, they did not process the opaque condition in the same way. The two native speaker groups did process opaque words similarly in terms of priming effect magnitude and this suggests native speakers are able to take advantage of the structural relationships between opaque word pairs, but non-native speakers are not.

While vocabulary proficiency did reveal differences in morphological processing in both the adult groups, it was not a useful individual difference measure for the masked priming task. English reading experience showed much stronger correlations with task measures for both learner groups. Also, the reading span task did not correlate with any dependent measure which shows reading span does not influence morphological processing strategies.

This research has implications for the classroom for both developing readers (children) and adult learners of English. For developing readers, understanding how complex words are decomposed lends insight insight into what parts of complex words are salient to children. While most literacy research recommends explicit instruction of all word segments including root (base) words and suffixes (Carlisle, 2007), results from experiment 3 suggest that children would especially benefit from explicit suffix instruction compared to base word instruction.

This research also shows stronger correlations between child reading level and task performance than vocabulary level. This means that independent reading is more important for developing morphological decomposition skill than vocabulary development. Children who struggle to generalize suffix rules should read more often either at home or in the classroom.

For adult learners of English, this research shows that explicit instruction of English suffixes would also be useful. As shown in the masked priming task, both learner groups are still learning what kinds of words can be broken down further. The word *hunter* can be separated into two parts but *castle* cannot. Priming effects showed both learner groups may be breaking down words like *castle* during visual word recognition. This means that explicit instruction is needed for both groups about which word endings are suffixes and which word endings are not suffixes.

For second language learners of English, practice or lack of practice with suffix decomposition in their native language may influence their skill in appropriate suffix decomposition in English.

Finally, much like irregular and regular verb endings are taught in class, explicit instruction regarding how suffixes relate to the base word may help clue learners into how an ending like -er is used in both transparent and opaque word types.

Study Limitations

For the English word targets, one study limitation is that the task did not produce effects of morphological processing in the learner groups. However, it is clear in both the nonword and masked priming tasks that both of the learner groups have knowledge of English morphological structure and are applying strategies using this information in a consistent way. This means that the English target words were not difficult enough for the readers or that the task needs to be designed in a way that provides clearer opportunities for morphological processing that is measurable.

Both the nonword targets and masked priming task produced clear effects across groups, even in the child group which had a small sample size. I would not change anything about either task and there were interesting results in both analyses to provide areas for future research.

In the population samples, one limitation to the generalizability of the results is that the native adult group consisted of highly educated college students which may explain why vocabulary effects were not evident in five of the six dependent measures for this study. For the child group, one limitation is that I was unable to recruit a larger population sample balanced by age and grade. This limited my study in two ways, first I was unable to find as many effects in each task analyses as I would have liked and second, I was unable to evaluate how individual differences in vocabulary and grade affect task performance as I was able for the adult groups. Therefore, I am unsure whether task effects, specifically priming effects are consistent across grade levels.

Future Research

Future research in this area should concentrate on two goals. First to create an effective online task that evaluates morphological processing of English words that can also serve as an effective baseline for learner groups. Even though the task in this study produced clear effects in the native adult group, I ultimately was unable to use this information to understand how learner groups would behave on the same task.

The second goal would be to understand the strategies being using during form and opaque priming in learner groups. Specifically, there needs to be further investigation on whether positive priming occurs in the form condition because of shared orthography or base word stripping. More information is needed about how learner groups use the suffix in the opaque condition. Rastle and Davis (2008) hypothesize that adult native speakers automatically decompose the suffix in the opaque condition in order to access the word more quickly. However, in both learner groups, this strategy is not utilized. Are learners ignoring the suffix or processing the suffix but also processing semantics inhibiting the response? The results from the current study cannot answer these questions.

In conclusion, variables relating to age, vocabulary proficiency and native language status influence how readers process English morphemes in complex English words and nonwords. Some strategies used by the readers are shared across age and learner group and there are also some clear differences. What is evident is that all of the readers are aware of these structures and are using them discriminately and deliberately to enable complex word processing.

Appendix A. Experiment 1. Language Questionnaire

Age:

Sex:

Major:

What languages do you speak fluently?

What languages have you had exposure to in the classroom?

Outside the classroom?

If you are not an English speaker, at what age did you learn English?

What percentage of your daily speech is in English?

During a school semester, how many hours do you spend each week reading for pleasure? For classes?

In a year, how many books do you read for pleasure?

Do you think your vocabulary knowledge is adequate for college courses?

Do you think your reading ability is adequate for college courses?

In your native language have you ever…

-had speech therapy?

-had a reading problem identified?

-had any other language problem identified?

Received a diagnosis of a learning difficulty?

Been tutored due to difficulty in school?

If yes in what areas?

Ever had problems with reading?

With spelling?

Problems in another language?

More than average difficulty learning a foreign language?

Does anyone in your family have a learning difficulty?

Appendix B. Experiment 2. Language Questionnaire (Li et al., 2014)

1. Participant ID number   2. Age   3. Gender   4. Education

5. Indicate your native language(s) and any other languages you have studied or learned, the age at which you started using each language in terms of listening, speaking, reading, and writing, and the total number of years you have spent using each language.
Language
Listening
Speaking
Reading
Writing
Years of Use*
*Notes: For Years of Use, you may have learned a language, stopped using it, and then started using it again. Please give the total number of years.

6. Country of Origin:

7. If you have lived or traveled in countries other than your country of residence for three months or more, then indicate the name of the country, your length of stay (in Months), the language you used, and the frequency of your use of the language for each country.
Country:
Length of stay (in Months)*:
Language:
Frequency of use:
*Note: You may have been to the country on multiple occasions, each for a different length of time. Add all the trips together

8. Rate your language learning skill. In other words, how good do you feel you are at learning new languages, relative to your friends or other people you know?
very poor to excellent (drop down box)

9. Rate your current ability in terms of listening, speaking, reading, and writing in each of the languages you have studied or learned (including the native language).
Language: very poor to excellent (drop down box)
Listening: very poor to excellent (drop down box)
Speaking: very poor to excellent (drop down box)
Reading: very poor to excellent (drop down box)
Writing: very poor to excellent (drop down box)

10. Use the comment box below to provide any other information about your language background or usage.

Appendix C. Child Reading History Questionnaire

Age:

Sex:

Grade:

What languages does your child speak fluently?

If your child is not a native English speaker, at what age did he or she learn English?

What percentage of your child's daily speech is in English?

How many hours each week does your child read or do you read to him or her?

Has your child ever…

-had speech therapy?

-had a reading problem identified?

-had any other language problem identified?

Received a diagnosis of a learning difficulty?

Been tutored due to difficulty in school?

If yes in what areas?

Does anyone in your family have a learning difficulty?

Appendix D. Adult English Vocabulary Test

Truculent
1. satisfied
2. brilliant
3. fawning
4. combative
5. reproaching

Discrete
1. prudent
2. judicious
3. stunted
4. stringent
5. separate

Consensus
1. poll
2. conference
3. agreement
4. attitude
5. honesty

Indigenous
1. elevating
2. destitute
3. livid
4. insulting
5. native

Proclivity
1. propensity
2. buoyancy
3. sacrifice
4. hegemony
5. illegibility

Sobriquet
1. debutante
2. nickname
3. puppet
4. caretaker
5. connection

Surrogate
1. will
2. substitute
3. criminal court
4. stratagem
5. duplicate

Aplomb
1. self-assurance
2. stodginess
3. sturdiness
4. overeager
5. narrow minded

Covert
1. envious
2. timid
3. protected
4. secret
5. smug

Unassuaged
1. unseen
2. unrelieved
3. unconvinced
4. unwashed
5. unrestricted

Fecund
1. fruitful
2. offensive
3. decaying
4. obliging
5. feverish

Arcane
1. curved
2. mysterious
3. risky
4. inventive
5. counterfeit

Flaunt
1. display brazenly
2. disobey insolently
3. question extensively
4. hit repeatedly
5. illuminate brightly

Moraine
1. hilly country
2. swamp
3. glacial deposit
4. small island
5. peninsula

Nebulous
1. false
2. basic
3. noxious
4. cloudy
5. noisy

Celerity
1. slipperiness
2. fame
3. speed
4. grace
5. sincerity

Bifurcate
1. fork
2. lie
3. translate
4. stagger
5. destroy

Limpid
1. deep
2. musical
3. still
4. decorative
5. clear

Slake
1. kill
2. quench
3. smear
4. freshen
5. imprison

Bellicose
1. resonant
2. beautiful
3. warlike
4. indulgent
5. overweight

Eschew
1. digest
2. despise
3. dry up
4. violate
5. avoid

Redolent
1. fragrant
2. lazy
3. difficult
4. suspicious
5. restless

Enervated
1. brash
2. excited
3. energetic
4. weak
5. encouraged

Euphoria
1. loss of memory
2. feeling of well being
3. loss of power of speech
4. feeling of loneliness
5. thoughts of suicide

Lachrymose
1. fearful
2. dark
3. curious
4. critical
5. tearful

Querulous
1. peevish
2. strange
3. questioning
4. exacting
5. poisonous

Ambulatory
1. able to walk
2. superior
3. related
4. imaginary
5. susceptible to disease

Adequate
1. eloquent
2. sufficient
3. water related
4. auspicious
5. unknown

Immerse
1. convert
2. erase
3. emphasize
4. reform
5. submerge

Plummet
1. ignore
2. punish
3. understand
4. scream
5. plunge

Notorious
1. capable
2. negotiable
3. infamous
4. secure
5. naive

Stagnant
1. vulnerable
2. doubtful
3. unchanging
4. difficult
5. dissimilar

Concise
1. terse
2. hungry
3. tired
4. painful
5. inventive

Jealous
1. talkative
2. truthful
3. jiggling
4. envious
5. avid

Chamber
1. pastry
2. goblet
3. room
4. echo
5. rocket

Plead
1. refuse
2. verify
3. scold
4. implore
5. discredit

Toil
1. labor
2. material
3. cleanse
4. bury
5. weep

Extravagant
1. unattractive
2. lavish
3. unobtrusive
4. grateful
5. miniscule

Humiliate
1. enjoy
2. act kindly
3. escape
4. praise
5. shame

Edifice
1. education
2. spokesperson
3. prophet
4. building
5. jargon

Random
1. new concept
2. fixed idea
3. without a method
4. updated
5. thoughtful

Narrate
1. to be quiet
2. a suggestion
3. a test
4. a greeting
5. a verbal account

Massive
1. convince
2. large
3. challenge
4. emphasize
5. evaluate

Inflate
1. fill
2. slide
3. wave
4. depend
5. protect

Misery
1. break
2. write
3. infect
4. distress
5. tiny

Install
1. to teach
2. to call
3. to shock
4. to shave
5. to place

Ration
1. a warning
2. wrap something
3. fixed amount
4. to report
5. strong willed

Rigid
1. effective
2. survived
3. productive
4. not flexible
5. carbonated

Span
1. range
2. inject
3. fail
4. allow
5. choose

Ruin
1. catch
2. offer
3. draw
4. destroy
5. improve

Appendix E. Lexical Decision Task (Casalis et al., 2015a)

W+ = base word present, S+ = suffix present

W+S+		W+S-		W-S+		W-S-	
Word	Nonword	Word	Nonword	Word	Nonword	Word	Nonword
bony	cabber	barrow	birdet	docile	cusky	apple	birtle
carriage	clockage	bullet	costle	crazy	demmy	arrow	bromel
cloudy	cutous	button	cupple	duty	dilly	borrow	cottle
cookery	farly	castle	cuttle	lobster	draly	bottle	cunon
daily	hidder	fellow	darlow	finger	erdious	elbow	forrow
digger	hurriage	freeze	dollet	glitter	foadous	follow	geavow
dirty	hurter	funnel	dullow	hoover	gopter	helmet	gleeze
driver	knower	handle	foodle	jealous	hettage	kettle	hannow
duster	lamper	jacket	gappow	leisure	hosper	lemon	olsow
failure	legger	needle	hillet	mirage	josy	meadow	pirdet
famous	lifty	pillow	kneeze	future	moacher	narrow	pittle
farmer	menny	rocket	landle	nasty	oasher	nugget	quabbel
gracious	nailure	saddle	lucket	obvious	paiture	quarrel	segget
hopper	peacher	settle	missow	leather	raby	sneeze	spetle
player	platious	shadow	notton	pony	reanage	staple	sullow
robber	prizer	single	rollow	pretty	sebber	stomach	trenach
sadly	proudy	spinach	runnel	sausage	slinny	target	ullow
sunny	ruly	wallet	shamow	sister	stanner	travel	uttle
teacher	sealy	window	songle	spider	stiner	turtle	volmet
wreckage	weekery	yellow	stirach	weather	teeker	velvet	walfet

Appendix F. Masked Priming Task (Li et al., 2017)

Pair Type= T-Transparent, O-Opaque, F-Form; Primes not covered by the mask are bolded

Pair type	Target	Related Prime	Unrelated Prime
T	sail	sailor	editor
T	proper	**properly**	trainer
T	farm	farmer	**eastern**
T	pain	**painful**	**speaker**
T	press	**pressure**	**careful**
T	weak	**weakness**	**harmful**
T	wealth	**wealthy**	yearly
T	peace	**peaceful**	writer
T	gold	golden	bossy
T	west	**western**	dentist
T	rain	rainy	endless
T	argue	**argument**	officer
T	lead	leader	joyous
T	own	ower	**sharply**
T	guilt	guilty	**hopeful**
T	near	nearly	neatly
T	wood	wooden	**quickly**
T	hunt	hunter	**arrival**
T	rare	rarely	**lightly**
T	soft	softly	**grateful**
O	tail	tailor	**closely**
O	tend	tender	scenic
O	treat	treaty	**shoulder**
O	secret	**secretary**	**relative**
O	slip	**slipper**	plenty
O	sweat	**sweater**	**freedom**
O	show	shower	**shortly**
O	fact	factory	**fearful**
O	custom	**customer**	hunger
O	form	formal	warmth
O	apart	**apartment**	actual
O	wonder	**wonderful**	**national**
O	hard	hardly	manner
O	arm	army	**thunder**
O	mess	**message**	**princess**
O	sum	summer	purely
O	corn	corner	sunny
O	flow	flower	**rapidly**
O	cent	center	**warfare**

O	part	party	linear
F	pill	pillow	**success**
F	pack	packet	**company**
F	pea	peach	common
F	thin	think	proud
F	urge	urgent	**feature**
F	rock	rocket	**biscuit**
F	opera	**operate**	**conduct**
F	tea	teach	punish
F	exam	**example**	casual
F	stand	**standard**	**finance**
F	ten	tense	**feather**
F	plan	planet	honor
F	since	**sincere**	**marriage**
F	wall	wallet	wisdom
F	prince	**principle**	import
F	harm	**harmony**	**several**
F	sea	season	active
F	mark	market	**passive**
F	sing	single	**satisfy**
F	elect	**electric**	**combine**

Mask Debriefing**.**

Subject #_______

During the task, did you observe anything on the screen right before you responded whether the word was and English word or not? If so, what?

If you saw something before the word, were you able to read it?

If you were able to read it, did you notice any patterns?

Appendix G. Mixed Effects Model Statistical Output for Experiments 1-3

The most complex model that would converge was used in each analysis and corresponding fixed factor coefficients are reported here. Any simple effects models entered after the initial model are reported after the complex models along with their coefficients. Following Li et al. (2017), I give the key for the factor labels at the beginning and then report the estimate, standard error, degrees of freedom, t-value or z-value and p-value for the model intercept and all fixed factors.

The theoretical guidelines used for simplifying non-converging models are discussed in the results section for experiment #1. The below convergence test was used after any convergence warning. Only models that passed the test (returned values less than .001) were reported in the results section.

```
# checking severity of convergence warnings
relgrad <- with(model@optinfo$derivs,solve(Hessian,gradient))
max(abs(relgrad))
```

This test was recommended in the below online conversation about convergence issues in the lme4 package. This test uses a slower and more precise estimation of model fit than the initial summary() function when retrieving model results.

https://github.com/lme4/lme4/issues/120

https://rstudio-pubs-static.s3.amazonaws.com/33653_57fc7b8e5d484c909b615d8633c01d51.html

Lexical Decision Task – English Words

Accuracy

score=accuracy, base=base word present/absent, suffix=suffix present/absent, vocabgroup=vocabulary group, high/low, readage= English reading acquisition group, early/late, word= target word, group = participant group

Response Times

react= response times, base=base word present/absent, suffix=suffix present/absent, vocabgroup=vocabulary group, high/low, readage= English reading acquisition group, early/late, word= target word, group = participant group

Native English Group

```
#accuracy models
score1.1 <-glmer(score ~ base * suffix * vocabgroup + (1|subject),
data=L1english, family=binomial)
```

```
## Fixed effects:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)              2.9005     0.2299  12.616  < 2e-16 ***
## base                     1.4718     0.5005   2.941 0.003273 **
## suffix                   1.5238     0.5484   2.779 0.005460 **
## vocabgroup               1.2128     0.3758   3.227 0.001250 **
## base:suffix             -2.7719     0.7539  -3.677 0.000236 ***
## base:vocabgroup         -1.3201     0.6748  -1.956 0.050431 .
## suffix:vocabgroup       -1.1355     0.7492  -1.516 0.129593
## base:suffix:vocabgroup   2.1248     1.0167   2.090 0.036635 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix vcbgrp bs:sff bs:vcb sffx:v
## base        -0.414
## suffix      -0.377  0.175
## vocabgroup  -0.557  0.255  0.233
## base:suffix  0.273 -0.664 -0.727 -0.169
## base:vcbgrp  0.307 -0.741 -0.129 -0.553  0.492
## sffx:vcbgrp  0.276 -0.128 -0.732 -0.498  0.532  0.277
## bs:sffx:vcb -0.202  0.492  0.539  0.367 -0.741 -0.663 -0.737
```

```
score1.2 <-glmer(score ~ base * suffix * vocabgroup + (1|word),
data=L1english, family=binomial)

## Fixed effects:
##                           Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 3.1414     0.3019  10.405  < 2e-16 ***
## base                        1.4991     0.5746   2.609  0.00909 **
## suffix                      1.5558     0.6237   2.494  0.01262 *
## vocabgroup                  1.2334     0.3777   3.266  0.00109 **
## base:suffix                -2.6927     0.8637  -3.118  0.00182 **
## base:vocabgroup            -1.3414     0.6780  -1.979  0.04786 *
## suffix:vocabgroup          -1.1557     0.7527  -1.535  0.12467
## base:suffix:vocabgroup      2.1781     1.0234   2.128  0.03332 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix vcbgrp bs:sff bs:vcb sffx:v
## base        -0.484
## suffix      -0.444  0.241
## vocabgroup  -0.425  0.226  0.208
## base:suffix  0.327 -0.664 -0.721 -0.150
## base:vcbgrp  0.236 -0.649 -0.116 -0.557  0.431
## sffx:vcbgrp  0.213 -0.113 -0.646 -0.502  0.466  0.279
## bs:sffx:vcb -0.154  0.430  0.476  0.369 -0.650 -0.662 -0.735
#accuracy simple effects
L1english.1 <- rename(filter(L1english,base==1))
score2 <-glmer(score ~ suffix + (1|subject) + (1|word), data=L1english.1,
family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.9011     0.4518  10.848   <2e-16 ***
## suffix       -0.5784     0.5050  -1.145    0.252
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.602
```

```
L1english.2 <- rename(filter(L1english,base==0))
score3 <-glmer(score ~ suffix + (1|subject) + (1|word), data=L1english.2,
family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.7184     0.2910  12.780   <2e-16 ***
## suffix        0.9893     0.4128   2.397   0.0165 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.394
L1english.3 <- rename(filter(L1english,vocabgroup==1))
score4 <-glmer(score ~ suffix * base + (1|subject), data=L1english.3,
family=binomial)

## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.180119   0.004043 1033.99   <2e-16 ***
## suffix       0.390646   0.004041   96.67   <2e-16 ***
## base         0.152484   0.004041   37.73   <2e-16 ***
## suffix:base -0.649901   0.004040 -160.85   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) suffix base
## suffix      0.000
## base        0.000  0.000
## suffix:base 0.000  0.000  0.000

score4.1 <-glmer(score ~ suffix * base + (1|word), data=L1english.3,
family=binomial)
 ## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.1079     0.3040  13.513   <2e-16 ***
## suffix        0.3882     0.5109   0.760    0.447
## base          0.1517     0.4529   0.335    0.738
## suffix:base  -0.6471     0.6829  -0.948    0.343
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Correlation of Fixed Effects:
##             (Intr) suffix base
## suffix      -0.595
## base        -0.671  0.400
## suffix:base  0.445 -0.748 -0.663
```

```
L1english.4 <- rename(filter(L1english,vocabgroup==0))
score5 <-glmer(score ~ suffix * base + (1|subject) + (1|word),
data=L1english.4, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.2386     0.3544   9.138  < 2e-16 ***
## suffix        1.5595     0.6521   2.392  0.01678 *
## base          1.5060     0.6032   2.497  0.01253 *
## suffix:base  -2.7047     0.9039  -2.992  0.00277 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) suffix base
## suffix      -0.410
## base        -0.445  0.264
## suffix:base  0.309 -0.719 -0.665
L1english.5 <- rename(filter(L1english,vocabgroup==1 & base==1))
score6 <-glmer(score ~ suffix + (1|subject) + (1|word), data=L1english.5,
family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.8919     0.5589   8.753   <2e-16 ***
## suffix       -0.2157     0.5414  -0.398     0.69
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.475
L1english.6 <- rename(filter(L1english,vocabgroup==1 & base==0))
score7 <-glmer(score ~ suffix + (1|subject) + (1|word), data=L1english.6,
family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.1079     0.3040   13.51   <2e-16 ***
## suffix        0.3882     0.5108    0.76    0.447
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.595
```

```
L1english.7 <- rename(filter(L1english,vocabgroup==0 & base==1))
score8 <-glmer(score ~ suffix + (1|subject) + (1|word), data=L1english.7,
family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)   5.1208     0.7346   6.970 3.16e-12 ***
## suffix       -1.1179     0.7460  -1.498    0.134    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.615
L1english.8 <- rename(filter(L1english,vocabgroup==0 & base==0))
score9 <-glmer(score ~ suffix + (1|subject) + (1|word), data=L1english.8,
family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)   3.0461     0.3050   9.989  < 2e-16 ***
## suffix        1.5253     0.5871   2.598  0.00938 ** 
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.372

# build an interaction model
english.IN = lmer(react ~ suffix * base * vocabgroup + (1 + suffix|subject) +
(1|word), data=L1englishRT, REML=FALSE)

##
## Fixed effects:
##                          Estimate Std. Error t value
## (Intercept)             1.539e-01  6.529e-04 235.784
## suffix                 -8.977e-06  6.084e-04  -0.015
## base                    4.634e-04  5.759e-04   0.805
## vocabgroup              1.320e-04  7.256e-04   0.182
## suffix:base            -9.771e-04  8.441e-04  -1.158
## suffix:vocabgroup       1.512e-04  4.836e-04   0.313
## base:vocabgroup         1.908e-04  4.514e-04   0.423
## suffix:base:vocabgroup -1.674e-05  6.614e-04  -0.025
##
```

```
## Correlation of Fixed Effects:
##              (Intr) suffix base   vcbgrp sffx:b sffx:v bs:vcb
## suffix      -0.436
## base        -0.434  0.465
## vocabgroup  -0.689  0.167  0.152
## suffix:base  0.296 -0.708 -0.682 -0.103
## sffx:vcbgrp  0.233 -0.496 -0.228 -0.334  0.342
## base:vcbgrp  0.215 -0.231 -0.489 -0.307  0.334  0.461
## sffx:bs:vcb -0.147  0.347  0.334  0.210 -0.488 -0.698 -0.683
```

```
Anova(english.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                             Chisq Df Pr(>Chisq)
## (Intercept)            55594.2876  1     <2e-16 ***
## suffix                     0.0002  1     0.9882
## base                       0.6473  1     0.4211
## vocabgroup                 0.0331  1     0.8556
## suffix:base                1.3399  1     0.2470
## suffix:vocabgroup          0.0977  1     0.7546
## base:vocabgroup            0.1787  1     0.6725
## suffix:base:vocabgroup     0.0006  1     0.9798
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Non-Native English Group

```
#accuracy models
score1 <-glmer(score ~ base * suffix + readage + (1|subject) + (1|word),
data=L2english, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.8278     0.4459   8.585   <2e-16 ***
## base          1.0605     0.5458   1.943   0.0520 .
## suffix        0.7479     0.5606   1.334   0.1821
## readage      -0.3322     0.3481  -0.954   0.3400
## base:suffix  -1.3769     0.8001  -1.721   0.0853 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix readag
## base        -0.507
## suffix      -0.486  0.413
## readage     -0.410 -0.002 -0.002
## base:suffix  0.337 -0.683 -0.701  0.002
```

```
score3 <-glmer(score ~ base * suffix * vocabgroup + (1 + base|subject),
data=L2english, family=binomial)
## Fixed effects:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)              2.4757     0.2447  10.117   <2e-16 ***
## base                     0.7341     0.3181   2.307   0.0210 *
## suffix                   0.3615     0.2828   1.278   0.2012
## vocabgroup               0.9038     0.3758   2.405   0.0162 *
## base:suffix             -0.8740     0.4289  -2.038   0.0416 *
## base:vocabgroup          0.7792     0.6376   1.222   0.2217
## suffix:vocabgroup        1.1365     0.6219   1.827   0.0676 .
## base:suffix:vocabgroup  -1.8037     0.9117  -1.978   0.0479 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix vcbgrp bs:sff bs:vcb sffx:v
## base        -0.530
## suffix      -0.429  0.331
## vocabgroup  -0.583  0.290  0.283
## base:suffix  0.283 -0.674 -0.660 -0.186
## base:vcbgrp  0.223 -0.439 -0.167 -0.424  0.337
## sffx:vcbgrp  0.200 -0.154 -0.455 -0.342  0.300  0.204
## bs:sffx:vcb -0.136  0.319  0.310  0.233 -0.471 -0.681 -0.683
score4 <-glmer(score ~ base * suffix * vocabgroup + (1|word), data=L2english,
family=binomial)

## Fixed effects:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)              2.9127     0.3710   7.850 4.15e-15 ***
## base                     0.7990     0.5489   1.456   0.1455
## suffix                   0.4291     0.5582   0.769   0.4420
## vocabgroup               0.8385     0.2913   2.878   0.0040 **
## base:suffix             -0.8772     0.8014  -1.095   0.2737
## base:vocabgroup          0.8372     0.6386   1.311   0.1899
## suffix:vocabgroup        1.1583     0.6309   1.836   0.0663 .
## base:suffix:vocabgroup  -1.8067     0.9263  -1.950   0.0511 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix vcbgrp bs:sff bs:vcb sffx:v
## base        -0.622
## suffix      -0.603  0.418
## vocabgroup  -0.280  0.193  0.191
## base:suffix  0.418 -0.686 -0.697 -0.133
## base:vcbgrp  0.131 -0.266 -0.086 -0.456  0.182
## sffx:vcbgrp  0.135 -0.089 -0.228 -0.461  0.159  0.211
## bs:sffx:vcb -0.092  0.184  0.155  0.314 -0.251 -0.689 -0.681
```

```
# build an RT interaction model
english.IN = lmer(react ~ suffix * base * readage + (1 + base|subject) +
(1|word), data=L2englishRT, REML=FALSE)
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)          0.1537320  0.0007724 199.043
## suffix               0.0005237  0.0006788   0.772
## base                 0.0005320  0.0006519   0.816
## readage             -0.0023012  0.0009643  -2.386
## suffix:base         -0.0014289  0.0009499  -1.504
## suffix:readage      -0.0006228  0.0005392  -1.155
## base:readage         0.0002642  0.0005228   0.505
## suffix:base:readage  0.0008780  0.0007516   1.168
##
## Correlation of Fixed Effects:
##             (Intr) suffix base   readag sffx:b sffx:r bs:rdg
## suffix      -0.394
## base        -0.328  0.467
## readage     -0.609  0.098  0.036
## suffix:base  0.282 -0.715 -0.679 -0.070
## suffix:redg  0.154 -0.383 -0.182 -0.255  0.274
## base:readag  0.056 -0.180 -0.389 -0.094  0.258  0.469
## sffx:bs:rdg -0.110  0.275  0.261  0.183 -0.383 -0.717 -0.673
```

```
Anova(english.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                          Chisq Df Pr(>Chisq)
## (Intercept)         39618.2585  1     < 2e-16 ***
## suffix                  0.5953  1     0.44038
## base                    0.6659  1     0.41447
## readage                 5.6949  1     0.01701 *
## suffix:base             2.2627  1     0.13252
## suffix:readage          1.3342  1     0.24805
## base:readage            0.2553  1     0.61334
## suffix:base:readage     1.3648  1     0.24271
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
english.IN2 = lmer(react ~ suffix * base * vocabgroup + (1 + base|subject) +
(1|word), data=L2englishRT, REML=FALSE)

## Fixed effects:
##                          Estimate Std. Error t value
## (Intercept)             1.511e-01  7.861e-04 192.213
## suffix                  1.259e-05  6.940e-04   0.018
## base                    8.831e-04  6.664e-04   1.325
## vocabgroup              2.833e-03  9.417e-04   3.008
## suffix:base            -6.524e-04  9.697e-04  -0.673
## suffix:vocabgroup       3.907e-04  5.419e-04   0.721
## base:vocabgroup        -3.990e-04  5.267e-04  -0.758
## suffix:base:vocabgroup -6.483e-04  7.552e-04  -0.859
##
## Correlation of Fixed Effects:
##             (Intr) suffix base   vcbgrp sffx:b sffx:v bs:vcb
## suffix      -0.396
## base        -0.323  0.468
## vocabgroup  -0.643  0.113  0.043
## suffix:base  0.284 -0.716 -0.679 -0.081
## sffx:vcbgrp  0.174 -0.430 -0.205 -0.262  0.308
## base:vcbgrp  0.065 -0.202 -0.434 -0.091  0.287  0.469
## sffx:bs:vcb -0.125  0.309  0.292  0.188 -0.427 -0.718 -0.671

Anova(english.IN2,type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                             Chisq Df Pr(>Chisq)
## (Intercept)            36945.9795  1  < 2.2e-16 ***
## suffix                     0.0003  1   0.985522
## base                       1.7561  1   0.185108
## vocabgroup                 9.0493  1   0.002628 **
## suffix:base                0.4526  1   0.501092
## suffix:vocabgroup          0.5198  1   0.470932
## base:vocabgroup            0.5740  1   0.448668
## suffix:base:vocabgroup     0.7371  1   0.390599
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Elementary Age Group

```
#accuracy model
score2 <-glmer(score ~ base * suffix + (1|subject) + (1|word), data=Eenglish,
family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.43817    0.42821   8.029 9.82e-16 ***
## base         0.12533    0.51905   0.241    0.809
## suffix       0.01281    0.54164   0.024    0.981
## base:suffix -0.21393    0.75536  -0.283    0.777
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix
## base        -0.609
## suffix      -0.579  0.471
## base:suffix  0.418 -0.687 -0.718

# build an interaction model
english.IN = lmer(react ~ suffix * base + (1 + base|subject) + (1|word),
data=EenglishRT, REML=FALSE)

## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  1.461e-01  6.412e-04 227.914
## suffix       8.177e-05  5.665e-04   0.144
## base         8.060e-04  5.447e-04   1.480
## suffix:base -3.147e-04  7.918e-04  -0.397
##
## Correlation of Fixed Effects:
##             (Intr) suffix base
## suffix      -0.393
## base        -0.465  0.463
## suffix:base  0.281 -0.715 -0.675
Anova(english.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                  Chisq Df Pr(>Chisq)
## (Intercept) 51944.8247  1     <2e-16 ***
## suffix          0.0208  1     0.8852
## base            2.1900  1     0.1389
## suffix:base     0.1580  1     0.6910
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

All Groups

```
#accuracy models
score2 <-glmer(score ~ base * suffix * group + (1|subject), data=allenglish,
family=binomial)
## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)         4.2129     0.3151  13.372  < 2e-16 ***
## base                1.2484     0.4638   2.692 0.007105 **
## suffix              1.4224     0.4806   2.960 0.003077 **
## group              -0.5629     0.1451  -3.878 0.000105 ***
## base:suffix        -2.5037     0.6836  -3.663 0.000249 ***
## base:group         -0.2468     0.1973  -1.251 0.210888
## suffix:group       -0.4223     0.2007  -2.104 0.035416 *
## base:suffix:group   0.6860     0.2887   2.376 0.017492 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix group  bs:sff bs:grp sffx:g
## base        -0.430
## suffix      -0.412  0.290
## group       -0.926  0.389  0.373
## base:suffix  0.286 -0.681 -0.706 -0.258
## base:group   0.422 -0.946 -0.285 -0.432  0.645
## suffix:grop  0.411 -0.290 -0.950 -0.423  0.671  0.322
## bs:sffx:grp -0.282  0.649  0.663  0.290 -0.948 -0.686 -0.698

score3 <-glmer(score ~ base * suffix * group + (1|word), data=allenglish,
family=binomial)

## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)         4.3366     0.3595  12.063  < 2e-16 ***
## base                1.4016     0.5846   2.398  0.01650 *
## suffix              1.6116     0.6003   2.685  0.00726 **
## group              -0.5218     0.1618  -3.225  0.00126 **
## base:suffix        -2.5946     0.8584  -3.023  0.00251 **
## base:group         -0.4032     0.2571  -1.569  0.11676
## suffix:group       -0.5058     0.2555  -1.979  0.04777 *
## base:suffix:group   0.7695     0.3704   2.077  0.03777 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Correlation of Fixed Effects:
##              (Intr) base   suffix group  bs:sff bs:grp sffx:g
## base         -0.600
## suffix       -0.586  0.366
## group        -0.834  0.513  0.500
## base:suffix   0.410 -0.682 -0.701 -0.351
## base:group    0.518 -0.864 -0.318 -0.630  0.590
## suffix:grop   0.525 -0.328 -0.852 -0.634  0.598  0.402
## bs:sffx:grp -0.362  0.601  0.590  0.438 -0.861 -0.695 -0.692

#test pair factor for unplanned comparisons
summary(m1<-aov(score~group, data=allenglish))
##                Df Sum Sq Mean Sq F value Pr(>F)
## group           2    5.1  2.5424   64.12 <2e-16 ***
## Residuals    9779  387.7  0.0396
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TukeyHSD(m1,"group",ordered=TRUE)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = score ~ group, data = allenglish)
##
## $group
##           diff        lwr        upr    p adj
## 2-3 0.03916690 0.02650265 0.05183115 0.00e+00
## 1-3 0.05990037 0.04750121 0.07229953 0.00e+00
## 1-2 0.02073347 0.01000193 0.03146501 1.79e-05

allE1 <- rename(filter(allenglish,base==1))
score.simple=glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allE1, family=binomial)
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.3313     0.2725  15.892   <2e-16 ***
## suffix       -0.4861     0.3418  -1.422    0.155
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.685
```

```
allE2 <- rename(filter(allenglish,base==0))
score.simple1=glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allE2, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.8489     0.2653  14.505   <2e-16 ***
## suffix        0.8023     0.4044   1.984   0.0473 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.521
allE3 <- rename(filter(allenglish,group==1))
score.simple2=glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allE3, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.2172     0.2688  15.688   <2e-16 ***
## suffix        0.2616     0.3982   0.657    0.511
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.578
allE4 <- rename(filter(allenglish,group==2))
score.simple3=glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allE4, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.33210    0.37246  11.631   <2e-16 ***
## suffix      -0.01866    0.45721  -0.041    0.967
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.604
```

```
allE5 <- rename(filter(allenglish,group==3))
score.simple4=glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allE5, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)   3.4343     0.3401  10.097   <2e-16 ***
## suffix        0.0146     0.4116   0.035    0.972    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.540
```

```
allE6 <- rename(filter(allenglish,group==1))
score.simple5=glmer(score ~ base * suffix + (1 + base|subject) + (1|word),
data=allE6, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)   3.8486     0.3142  12.250   <2e-16 ***
## base          0.9185     0.4869   1.886   0.0592 .  
## suffix        1.0140     0.4702   2.157   0.0310 *  
## base:suffix  -1.6146     0.6491  -2.488   0.0129 *  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix
## base        -0.576              
## suffix      -0.457  0.307       
## base:suffix  0.340 -0.583 -0.722
allE7 <- rename(filter(allenglish,group==2))
score.simple6=glmer(score ~ base * suffix + (1 + base|subject) + (1|word),
data=allE7, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)   3.8698     0.4207   9.199   <2e-16 ***
## base          0.7946     0.5587   1.422   0.1550    
## suffix        0.6807     0.5429   1.254   0.2098    
## base:suffix  -1.3346     0.7722  -1.728   0.0839 .  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Correlation of Fixed Effects:
##             (Intr) base   suffix
## base        -0.601              
## suffix      -0.504  0.389       
## base:suffix  0.348 -0.646 -0.704
```

```
allE8 <- rename(filter(allenglish,group==3))
score.simple7=glmer(score ~ base * suffix + (1 + base|subject) + (1|word),
data=allE8, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)  3.49592    0.44519   7.853 4.08e-15 ***
## base         0.01827    0.54516   0.034    0.973    
## suffix       0.01404    0.54570   0.026    0.979    
## base:suffix -0.21108    0.75980  -0.278    0.781    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix
## base        -0.644              
## suffix      -0.561  0.451       
## base:suffix  0.406 -0.659 -0.719
```

```
allE9 <- rename(filter(allenglish,group==1 & base==1))
score.simple8=glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allE9, family=binomial)
```

```
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)   4.8467     0.6320   7.669 1.74e-14 ***
## suffix       -0.4230     0.7494  -0.564    0.572    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.765
```

```
allE10 <- rename(filter(allenglish,group==1 & base==0))
score.simple9=glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allE10, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)   3.6541     0.2646  13.811   <2e-16 ***
## suffix        1.5670     0.7714   2.031   0.0422 *  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.278
```

```
# build an interaction model
english.IN = lmer(react ~ suffix * base * group + (1 + suffix|subject) +
(1|word), data=allenglishRT, REML=FALSE)
## Fixed effects:
##                      Estimate Std. Error t value
## (Intercept)        1.586e-01  8.380e-04 189.246
## suffix            -5.597e-05  6.758e-04  -0.083
## base               4.823e-04  6.537e-04   0.738
## group             -3.727e-03  4.254e-04  -8.761
## suffix:base       -1.394e-03  9.518e-04  -1.465
## suffix:group       1.389e-04  3.289e-04   0.422
## base:group         1.111e-04  3.212e-04   0.346
## suffix:base:group  2.770e-04  4.654e-04   0.595
##
## Correlation of Fixed Effects:
##             (Intr) suffix base   group  sffx:b sffx:g bs:grp
## suffix      -0.366
## base        -0.382  0.475
## group       -0.898  0.297  0.310
## suffix:base  0.267 -0.710 -0.687 -0.218
## suffix:grop  0.310 -0.829 -0.402 -0.360  0.588
## base:group   0.320 -0.398 -0.836 -0.370  0.574  0.480
## sffx:bs:grp -0.226  0.586  0.577  0.261 -0.832 -0.707 -0.690
```

```
Anova(english.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                        Chisq Df Pr(>Chisq)
## (Intercept)       35814.0429  1     <2e-16 ***
## suffix                0.0069  1     0.9340
## base                  0.5444  1     0.4606
## group                76.7484  1     <2e-16 ***
## suffix:base           2.1453  1     0.1430
## suffix:group          0.1784  1     0.6728
## base:group            0.1196  1     0.7295
## suffix:base:group     0.3543  1     0.5517
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#test pair factor for unplanned comparisons
summary(m1<-aov(react~group, data=allenglishRT))
##                Df Sum Sq Mean Sq F value Pr(>F)
## group           2 0.0811 0.04053    1089 <2e-16 ***
## Residuals    9239 0.3440 0.00004
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
TukeyHSD(m1,"group",ordered=TRUE)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = react ~ group, data = allenglishRT)
##
## $group
##             diff          lwr         upr p adj
## 2-3 0.0066897468 0.0062840126 0.007095481     0
## 1-3 0.0075981811 0.0072021213 0.007994241     0
## 1-2 0.0009084343 0.0005723688 0.001244500     0
```

Lexical Decision Task – Nonwords

Accuracy

score=accuracy, base=base word present/absent, suffix=suffix present/absent, vocabgroup=vocabulary group, high/low , readage= English reading acquisition group, early/late, word= target word, group = participant group

Response Times

react= response times, base=base word present/absent, suffix=suffix present/absent, vocabgroup=vocabulary group, high/low, readage= English reading acquisition group, early/late, word= target word, group = participant group

Native English Group

```
#accuracy models
nonword.scoreIN <-glmer(score ~ base * suffix + vocabgroup + (1 +
suffix|subject) + (1|word), data=L1nonword, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.8944     0.3274   8.841  < 2e-16 ***
## base         -0.2339     0.3633  -0.644  0.51968
## suffix        0.1222     0.3788   0.323  0.74699
## vocabgroup    0.4063     0.2478   1.639  0.10114
## base:suffix  -1.5755     0.5062  -3.113  0.00185 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Correlation of Fixed Effects:
##             (Intr) base   suffix vcbgrp
## base        -0.573
## suffix      -0.590  0.492
## vocabgroup  -0.476  0.000  0.033
## base:suffix  0.412 -0.717 -0.713 -0.010

#accuracy simple effects
L1nonword.1 <- rename(filter(L1nonword,base==1))
score1 <-glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=L1nonword.1, family=binomial)
 ## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.9629     0.2998    9.883  < 2e-16 ***
## suffix       -1.5221     0.3710   -4.103 4.07e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.703

L1nonword.2 <- rename(filter(L1nonword,base==0))
score2 <-glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=L1nonword.2, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.0478     0.2861  10.653   <2e-16 ***
## suffix        0.3090     0.3938   0.785    0.433
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.594

# build RT interaction model
nonword.IN1 = lmer(react ~ suffix * base * vocabgroup + (1 +
base*suffix|subject) + (1|word), data=L1nonwordRT, REML=FALSE)
## Fixed effects:
##                          Estimate Std. Error t value
## (Intercept)             1.503e-01  7.747e-04 194.002
## suffix                 -1.157e-04  5.505e-04  -0.210
## base                   -1.216e-03  5.673e-04  -2.144
## vocabgroup              3.957e-04  9.064e-04   0.437
## suffix:base            -2.272e-03  7.938e-04  -2.862
## suffix:vocabgroup       8.845e-05  4.385e-04   0.202
## base:vocabgroup         5.946e-04  4.697e-04   1.266
## suffix:base:vocabgroup  4.352e-04  6.470e-04   0.673
##
```

```
## Correlation of Fixed Effects:
##              (Intr) suffix base   vcbgrp sffx:b sffx:v bs:vcb
## suffix       -0.360
## base         -0.334  0.510
## vocabgroup   -0.724  0.124  0.107
## suffix:base   0.147 -0.691 -0.691  0.002
## sffx:vcbgrp   0.182 -0.498 -0.272 -0.249  0.342
## base:vcbgrp   0.152 -0.262 -0.521 -0.207  0.343  0.527
## sffx:bs:vcb   0.003  0.334  0.348 -0.006 -0.512 -0.671 -0.668
```

```
Anova(nonword.IN1, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                             Chisq Df Pr(>Chisq)
## (Intercept)             37636.8223  1  < 2.2e-16 ***
## suffix                     0.0441  1   0.833600
## base                       4.5957  1   0.032052 *
## vocabgroup                 0.1905  1   0.662461
## suffix:base                8.1934  1   0.004204 **
## suffix:vocabgroup          0.0407  1   0.840132
## base:vocabgroup            1.6025  1   0.205549
## suffix:base:vocabgroup     0.4525  1   0.501150
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#RT simple effects
L1nonword1 <- rename(filter(L1nonwordRT,base==1))
simple1=lmer(react ~ suffix + (1 + suffix|subject) + (1|word),
data=L1nonword1, REML=FALSE)
```

```
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.1496838  0.0005686 263.254
## suffix      -0.0020766  0.0005331  -3.895
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.519
## convergence code: 0
## boundary (singular) fit: see ?isSingular
Anova(simple1)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## suffix 15.174  1  9.804e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
L1nonword2 <- rename(filter(L1nonwordRT,base==0))
simple2=lmer(react ~ suffix + (1 + suffix|subject) + (1|word),
data=L1nonword2, REML=FALSE)
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  1.506e-01  5.161e-04 291.740
## suffix      -6.123e-05  4.321e-04  -0.142
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.430
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
Anova(simple2)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## suffix 0.0201  1     0.8873
```

Non-native English Group

```
#accuracy models
non.scoreIN <-glmer(score ~ base * suffix * readage + (1|subject) + (1|word),
data=L2nonword, family=binomial)
## Fixed effects:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)           2.7165     0.3996   6.798 1.06e-11 ***
## base                 -0.1112     0.3504  -0.317   0.7510
## suffix                0.3477     0.3583   0.971   0.3318
## readage              -0.9294     0.4837  -1.922   0.0547 .
## base:suffix          -1.0918     0.5043  -2.165   0.0304 *
## base:readage         -0.5111     0.2834  -1.804   0.0713 .
## suffix:readage       -0.5184     0.2959  -1.752   0.0797 .
## base:suffix:readage   0.3599     0.4025   0.894   0.3712
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix readag bs:sff bs:rdg sffx:r
## base        -0.444
## suffix      -0.430  0.494
## readage     -0.622  0.140  0.133
## base:suffix  0.300 -0.695 -0.711 -0.091
## base:readag  0.206 -0.467 -0.232 -0.308  0.325
## suffix:redg  0.193 -0.227 -0.489 -0.290  0.349  0.502
## bs:sffx:rdg -0.138  0.329  0.361  0.207 -0.471 -0.701 -0.736
```

```
non.scoreIN2 <-glmer(score ~ base * suffix * vocabgroup + (1|subject) +
(1|word), data=L2nonword, family=binomial)
```

```
## Fixed effects:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)              1.9089     0.4061   4.701 2.59e-06 ***
## base                    -0.7398     0.3347  -2.210  0.02708 *
## suffix                  -0.2546     0.3366  -0.756  0.44957
## vocabgroup               0.6386     0.4903   1.303  0.19274
## base:suffix             -0.4278     0.4788  -0.894  0.37153
## base:vocabgroup          0.7586     0.2833   2.677  0.00742 **
## suffix:vocabgroup        0.7023     0.2956   2.376  0.01749 *
## base:suffix:vocabgroup  -1.0458     0.4017  -2.604  0.00923 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix vcbgrp bs:sff bs:vcb sffx:v
## base        -0.430
## suffix      -0.423  0.512
## vocabgroup  -0.629  0.119  0.116
## base:suffix  0.293 -0.696 -0.702 -0.079
## base:vcbgrp  0.174 -0.377 -0.203 -0.299  0.259
## sffx:vcbgrp  0.162 -0.195 -0.370 -0.281  0.260  0.491
## bs:sffx:vcb -0.116  0.263  0.272  0.201 -0.367 -0.702 -0.737
```

```
#accuracy simple effects
L2nonword.1 <- rename(filter(L2nonword,base==0))
non.score1 <-glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=L2nonword.1, family=binomial)
```

```
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.1610     0.3009   7.183 6.82e-13 ***
## suffix        0.4134     0.3378   1.224    0.221
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.301
```

```
L2nonword.2 <- rename(filter(L2nonword,base==1))
non.score2 <-glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=L2nonword.2, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.9054     0.3418   5.574 2.49e-08 ***
## suffix       -0.8440     0.3393  -2.487   0.0129 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.483
## convergence code: 0
## boundary (singular) fit: see ?isSingular

L2nonword.3 <- rename(filter(L2nonword,vocabgroup==1))
non.score3 <-glmer(score ~ base * suffix + (1 + base|subject) + (1|word),
data=L2nonword.3, family=binomial)
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.55800    0.35719   7.161 7.99e-13 ***
## base        -0.09156    0.37306  -0.245  0.80613
## suffix       0.43127    0.36679   1.176  0.23969
## base:suffix -1.39914    0.51176  -2.734  0.00626 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix
## base        -0.562
## suffix      -0.487  0.464
## base:suffix  0.346 -0.668 -0.716
## convergence code: 0
## boundary (singular) fit: see ?isSingular

L2nonword.4 <- rename(filter(L2nonword,vocabgroup==0))
non.score4 <-glmer(score ~ base * suffix + (1 + base|subject) + (1|word),
data=L2nonword.4, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.9314     0.4698   4.111 3.94e-05 ***
## base         -0.7069     0.3488  -2.026   0.0427 *
## suffix       -0.2464     0.3369  -0.731   0.4646
## base:suffix  -0.4623     0.4802  -0.963   0.3356
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Correlation of Fixed Effects:
##             (Intr) base   suffix
## base        -0.363
## suffix      -0.369  0.496
## base:suffix  0.258 -0.681 -0.701
```

```
L2nonword.5 <- rename(filter(L2nonword,vocabgroup==1))
non.score5 <-glmer(score ~ base + suffix + (1 + base|subject) + (1|word),
data=L2nonword.5, family=binomial)
```

```
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.9469     0.3468   8.497   <2e-16 ***
## base         -0.7913     0.2924  -2.707   0.0068 **
## suffix       -0.2964     0.2717  -1.091   0.2753
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr) base
## base   -0.491
## suffix -0.389  0.005
```

```
L2nonword.6 <- rename(filter(L2nonword,vocabgroup==0))
non.score6 <-glmer(score ~ base + suffix + (1 + base|subject) + (1|word),
data=L2nonword.6, family=binomial)
```

```
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.0521     0.4559   4.501 6.75e-06 ***
## base         -0.9374     0.2572  -3.644 0.000268 ***
## suffix       -0.4752     0.2420  -1.963 0.049613 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr) base
## base   -0.274
## suffix -0.275  0.040
```

```
L2nonword.7 <- rename(filter(L2nonword,vocabgroup==1 & base==0))
non.score7 <-glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=L2nonword.7, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.4802     0.3585   6.917  4.6e-12 ***
## suffix        0.9141     0.5167   1.769   0.0769 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.299
```

```
L2nonword.8 <- rename(filter(L2nonword,vocabgroup==1 & base==1))
non.score8 <-glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=L2nonword.8, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.3220     0.3099   7.494 6.71e-14 ***
## suffix       -0.7770     0.3690  -2.105   0.0353 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.449
```

```
# build an RT model
nonword.IN = lmer(react ~ suffix * base * readage + (1 + base|subject) +
(1|word), data=L2nonwordRT, REML=FALSE)
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)          1.498e-01  1.095e-03 136.865
## suffix              -8.977e-05  7.395e-04  -0.121
## base                -1.422e-03  7.602e-04  -1.871
## readage             -4.243e-03  1.481e-03  -2.864
## suffix:base         -1.052e-03  1.080e-03  -0.974
## suffix:readage      -5.039e-04  7.684e-04  -0.656
## base:readage        -2.548e-04  8.183e-04  -0.311
## suffix:base:readage  2.217e-04  1.173e-03   0.189
## Correlation of Fixed Effects:
##             (Intr) suffix base   readag sffx:b sffx:r bs:rdg
## suffix      -0.343
## base        -0.159  0.493
## readage     -0.652  0.124 -0.008
## suffix:base  0.235 -0.684 -0.675 -0.085
## suffix:redg  0.162 -0.466 -0.233 -0.263  0.319
## base:readag -0.010 -0.225 -0.474  0.000  0.307  0.477
## sffx:bs:rdg -0.106  0.305  0.305  0.174 -0.454 -0.656 -0.646
```

```
Anova(nonword.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                          Chisq Df Pr(>Chisq)
## (Intercept)         18731.9400  1  < 2.2e-16 ***
## suffix                  0.0147  1   0.903382
## base                    3.4992  1   0.061400 .
## readage                 8.2053  1   0.004177 **
## suffix:base             0.9485  1   0.330109
## suffix:readage          0.4301  1   0.511955
## base:readage            0.0970  1   0.755476
## suffix:base:readage     0.0357  1   0.850043
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
nonword.IN1 = lmer(react ~ suffix * base * vocabgroup + (1 + base|subject) +
(1|word), data=L2nonwordRT, REML=FALSE)
## Fixed effects:
##                          Estimate Std. Error t value
## (Intercept)             1.459e-01  1.201e-03 121.474
## suffix                 -4.436e-04  7.940e-04  -0.559
## base                   -1.197e-03  8.280e-04  -1.446
## vocabgroup              3.431e-03  1.551e-03   2.212
## suffix:base            -9.892e-04  1.185e-03  -0.834
## suffix:vocabgroup       2.102e-04  7.741e-04   0.272
## base:vocabgroup        -5.925e-04  8.301e-04  -0.714
## suffix:base:vocabgroup  6.959e-05  1.174e-03   0.059
##
## Correlation of Fixed Effects:
##             (Intr) suffix base   vcbgrp sffx:b sffx:v bs:vcb
## suffix      -0.330
## base        -0.130  0.479
## vocabgroup  -0.699  0.141 -0.009
## suffix:base  0.222 -0.669 -0.665 -0.095
## sffx:vcbgrp  0.186 -0.567 -0.271 -0.252  0.379
## base:vcbgrp -0.012 -0.264 -0.586  0.033  0.376  0.472
## sffx:bs:vcb -0.124  0.373  0.381  0.168 -0.585 -0.659 -0.646
```

```
Anova(nonword.IN1, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                             Chisq Df Pr(>Chisq)
## (Intercept)            14755.8922  1     < 2e-16 ***
## suffix                     0.3122  1     0.57636
## base                       2.0899  1     0.14828
## vocabgroup                 4.8934  1     0.02696 *
## suffix:base                0.6963  1     0.40403
## suffix:vocabgroup          0.0738  1     0.78594
## base:vocabgroup            0.5095  1     0.47534
## suffix:base:vocabgroup     0.0035  1     0.95274
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Elementary Age Group

```
#accuracy models
non.scoreIN <-glmer(score ~ base * suffix + (1 + base|subject) + (1|word),
data=Enonword, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.3232     0.3441   6.752 1.46e-11 ***
## base         -0.6853     0.3336  -2.055   0.0399 *
## suffix        0.1044     0.3116   0.335   0.7376
## base:suffix  -0.8298     0.4362  -1.902   0.0572 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix
## base        -0.690
## suffix      -0.447  0.461
## base:suffix  0.318 -0.636 -0.714

# build an interaction model
nonword.IN = lmer(react ~ suffix * base + (1 + base|subject) + (1|word),
data=EnonwordRT, REML=FALSE)

## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  0.1440277  0.0006667 216.023
## suffix      -0.0005251  0.0004973  -1.056
## base        -0.0014320  0.0005008  -2.859
## suffix:base -0.0005681  0.0007323  -0.776
##
```

```
## Correlation of Fixed Effects:
##             (Intr) suffix base
## suffix      -0.372
## base        -0.375  0.495
## suffix:base  0.253 -0.679 -0.684
Anova(nonword.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                  Chisq Df Pr(>Chisq)
## (Intercept) 46666.1100  1  < 2.2e-16 ***
## suffix          1.1153  1   0.290934
## base            8.1745  1   0.004248 **
## suffix:base     0.6019  1   0.437835
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

All Groups

```
#accuracy models
nonword.scoreIN <-glmer(score ~ base * suffix * group + (1 + suffix|subject)
+ (1|word), data=allnonword, family=binomial)
## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)        3.38100    0.38877   8.697  < 2e-16 ***
## base              -0.16935    0.38064  -0.445  0.65638
## suffix             0.28018    0.40775   0.687  0.49199
## group             -0.44995    0.18967  -2.372  0.01768 *
## base:suffix       -1.62100    0.53566  -3.026  0.00248 **
## base:group        -0.07320    0.17545  -0.417  0.67654
## suffix:group      -0.07399    0.18797  -0.394  0.69387
## base:suffix:group  0.21419    0.24887   0.861  0.38942
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) base   suffix group  bs:sff bs:grp sffx:g
## base        -0.513
## suffix      -0.483  0.489
## group       -0.876  0.400  0.371
## base:suffix  0.362 -0.709 -0.717 -0.281
## base:group   0.422 -0.816 -0.403 -0.487  0.578
## suffix:grop  0.391 -0.404 -0.831 -0.446  0.594  0.493
## bs:sffx:grp -0.300  0.574  0.586  0.345 -0.813 -0.703 -0.714
```

```
#accuracy simple effects
allnonword.1 <- rename(filter(allnonword,base==1))
score1 <-glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allnonword.1, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.3354     0.2007  11.637  < 2e-16 ***
## suffix       -1.1398     0.2404  -4.742 2.12e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.628

allnonword.2 <- rename(filter(allnonword,base==0))
score2 <-glmer(score ~ suffix + (1 + suffix|subject) + (1|word),
data=allnonword.2, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.5525     0.1833  13.928   <2e-16 ***
## suffix        0.3755     0.2277   1.649   0.0991 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.482

#test pair factor for unplanned comparisons
summary(m1<-aov(score~group, data=allnonword))
##                Df Sum Sq Mean Sq F value Pr(>F)
## group           2   21.8  10.898   77.02 <2e-16 ***
## Residuals   10479 1482.8   0.142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
TukeyHSD(m1,"group",ordered=TRUE)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = score ~ group, data = allnonword)
##
## $group
##            diff         lwr        upr     p adj
## 3-2 0.008029287 -0.01526879 0.03132737 0.6981732
## 1-2 0.095251126  0.07568642 0.11481583 0.0000000
## 1-3 0.087221839  0.06456273 0.10988095 0.0000000
```

```
# build an interaction model
nonword.IN = lmer(react ~ suffix * base * group + (1 + suffix|subject) +
(1|word), data=allnonwordRT, REML=FALSE)
## Fixed effects:
##                     Estimate Std. Error t value
## (Intercept)        0.1539599  0.0010385 148.247
## suffix             0.0002881  0.0006759   0.426
## base              -0.0006057  0.0006828  -0.887
## group             -0.0031856  0.0005289  -6.023
## suffix:base       -0.0026863  0.0009872  -2.721
## suffix:group      -0.0002956  0.0003375  -0.876
## base:group        -0.0002896  0.0003418  -0.847
## suffix:base:group  0.0007105  0.0004997   1.422
##
## Correlation of Fixed Effects:
##              (Intr) suffix  base  group  sffx:b sffx:g bs:grp
## suffix      -0.345
## base        -0.325  0.497
## group       -0.907  0.291  0.273
## suffix:base  0.221 -0.684 -0.691 -0.184
## suffix:grop  0.296 -0.856 -0.425 -0.338  0.585
## base:group   0.277 -0.424 -0.857 -0.317  0.592  0.494
## sffx:bs:grp -0.186  0.578  0.586  0.212 -0.859 -0.675 -0.683

Anova(nonword.IN)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##                     Chisq Df Pr(>Chisq)
## suffix            13.4925  1  0.0002395 ***
## base              51.7227  1  6.392e-13 ***
## group             46.8060  1  7.837e-12 ***
## suffix:base        8.5774  1  0.0034036 **
## suffix:group       0.0129  1  0.9096299
## base:group         0.0288  1  0.8653132
## suffix:base:group  2.0214  1  0.1550922
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#RT simple effects
allnonword1 <- rename(filter(allnonwordRT,base==1))
nonword.simple=lmer(react ~ suffix + (1 + suffix|subject) + (1|word),
data=allnonword1, REML=FALSE)

## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)  0.1471447  0.0005150 285.706
## suffix      -0.0017495  0.0004411  -3.966
##
```

```
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.356
Anova(nonword.simple)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## suffix 15.731  1  7.302e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

allnonword2 <- rename(filter(allnonwordRT,base==0))
nonword.simple1=lmer(react ~ suffix + (1 + suffix|subject) + (1|word),
data=allnonword2, REML=FALSE)
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  0.1482570  0.0004722  313.95
## suffix      -0.0001985  0.0003053   -0.65
##
## Correlation of Fixed Effects:
##        (Intr)
## suffix -0.389
Anova(nonword.simple1)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## suffix 0.4227  1     0.5156
summary(m1<-aov(react~group, data=allnonwordRT))
##               Df Sum Sq  Mean Sq F value Pr(>F)
## group          2 0.0553 0.027643   600.6 <2e-16 ***
## Residuals   8360 0.3848 0.000046
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
TukeyHSD(m1,"group",ordered=TRUE)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = react ~ group, data = allnonwordRT)
##
## $group
##            diff         lwr         upr p adj
## 2-3 0.005118499 0.004638098 0.005598900     0
## 1-3 0.006733378 0.006275520 0.007191235     0
## 1-2 0.001614879 0.001219183 0.002010574     0
```

Masked Priming Task

Accuracy

score=accuracy, relate=relatedness, pair = condition, vocabgroup: vocabulary group, high/low, readage= English reading acquisition group, early/late, word= target word, group = participant group

Response Times

react= response times, relate=relatedness, pair = condition, vocabgroup: vocabulary group, readage= English reading acquisition group, early/late, word= target word, group = participant group

Native English Group

```
#accuracy models
mask.scoreIN <-glmer(score~relate * pair + vocabgroup + (1|subject) +
(1|word), data=L1mask, family=binomial)

## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.55659    0.43580   8.161 3.32e-16 ***
## relate      -0.20670    0.55979  -0.369    0.712
## pair         0.06885    0.18568   0.371    0.711
## vocabgroup   0.21462    0.26241   0.818    0.413
## relate:pair  0.15902    0.27033   0.588    0.556
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Correlation of Fixed Effects:
##             (Intr) relate pair   vcbgrp
## relate      -0.636
## pair        -0.834  0.649
## vocabgroup  -0.341  0.000  0.000
## relate:pair  0.573 -0.920 -0.687  0.000

#enter the model
mask.IN=lmer(react ~ relate * pair * vocabgroup + (1 + relate*pair|subject) +
(1+relate|word), data=L1maskRT, REML=FALSE)## Fixed effects:
##                          Estimate Std. Error t value
## (Intercept)              1.402e-01  2.896e-04 484.247
## relate                  -4.724e-04  2.770e-04  -1.706
## pair                    -4.278e-05  8.967e-05  -0.477
## vocabgroup              -1.795e-04  3.621e-04  -0.496
## relate:pair              4.805e-04  1.258e-04   3.819
## relate:vocabgroup        1.486e-04  3.518e-04   0.422
## pair:vocabgroup          9.891e-05  1.166e-04   0.848
## relate:pair:vocabgroup -1.712e-04  1.636e-04  -1.046
##
```

```
## Correlation of Fixed Effects:
##              (Intr) relate pair   vcbgrp rlt:pr rlt:vc pr:vcb
## relate       -0.334
## pair         -0.595  0.627
## vocabgroup   -0.739  0.315  0.475
## relate:pair   0.375 -0.897 -0.666 -0.300
## relt:vcbgrp   0.310 -0.751 -0.494 -0.419  0.706
## pair:vcbgrp   0.457 -0.482 -0.769 -0.618  0.512  0.641
## rlt:pr:vcbg  -0.289  0.690  0.512  0.390 -0.769 -0.919 -0.666

Anova(mask.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                             Chisq Df Pr(>Chisq)
## (Intercept)            2.3450e+05  1  < 2.2e-16 ***
## relate                 2.9089e+00  1  0.0880938 .
## pair                   2.2760e-01  1  0.6333161
## vocabgroup             2.4570e-01  1  0.6201056
## relate:pair            1.4582e+01  1  0.0001342 ***
## relate:vocabgroup      1.7850e-01  1  0.6726900
## pair:vocabgroup        7.1950e-01  1  0.3963078
## relate:pair:vocabgroup 1.0940e+00  1  0.2955923
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#test simple effects in interaction
maskL1.1 <- rename(filter(L1maskRT,pair==1))
mask.simple=lmer(react ~ relate + (1 + relate|subject) + (1+relate|word),
data=maskL1.1, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 1.401e-01  2.019e-04 694.222
## relate      7.547e-05  1.233e-04   0.612
##
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.048

Anova(mask.simple)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 0.3747  1     0.5404
```

```
maskL1.2 <- rename(filter(L1maskRT,pair==2))
mask.simple1=lmer(react ~ relate + (1 + relate|subject) + (1 + relate|word),
data=maskL1.2, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 0.1401027  0.0001776 788.761
## relate      0.0002138  0.0001359   1.573
##
## Correlation of Fixed Effects:
##        (Intr)
## relate 0.040

Anova(mask.simple1)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 2.4735  1     0.1158

maskL1.3 <- rename(filter(L1maskRT,pair==3))
mask.simple2=lmer(react ~ relate + (1 + relate|subject) + (1 + relate|word),
data=maskL1.3, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 0.1401804  0.0001683 832.758
## relate      0.0008258  0.0001418   5.823
##
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.050

Anova(mask.simple2)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 33.904  1   5.79e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Non-native English Group

```
#test interaction model
mask.scoreIN <-glmer(score ~ relate * pair + readage + (1|subject) + (1 +
relate|word), data=L2mask, family=binomial)
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.2999     0.5165   6.389 1.68e-10 ***
## relate        0.5993     0.6315   0.949   0.3426
## pair          0.4173     0.2141   1.949   0.0513 .
## readage      -0.6941     0.2832  -2.451   0.0143 *
## relate:pair  -0.3193     0.3044  -1.049   0.2942
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) relate pair   readag
## relate       -0.589
## pair         -0.721  0.593
## readage      -0.352 -0.004 -0.007
## relate:pair   0.508 -0.889 -0.704  0.004

mask.score1 <-glmer(score ~ relate * pair + vocabgroup + (1|subject) + (1 +
relate|word), data=L2mask, family=binomial)
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   2.5414     0.4924   5.162 2.45e-07 ***
## relate        0.5992     0.6345   0.944  0.34496
## pair          0.4175     0.2152   1.939  0.05244 .
## vocabgroup    0.8365     0.2812   2.975  0.00293 **
## relate:pair  -0.3194     0.3058  -1.045  0.29621
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
##              (Intr) relate pair   vcbgrp
## relate       -0.623
## pair         -0.764  0.593
## vocabgroup   -0.186  0.004  0.008
## relate:pair   0.538 -0.889 -0.704 -0.004

#enter the model
mask.model=lmer(react ~ relate + pair + readage + (1 + pair|subject) +
(1|word), data=L2maskRT, REML=FALSE)
## Fixed effects:
##                Estimate Std. Error t value
## (Intercept)  1.400e-01  3.104e-04 451.042
## relate       3.394e-04  8.423e-05   4.030
## pair         1.108e-04  5.331e-05   2.079
## readage     -1.136e-03  3.629e-04  -3.130
```

```
## Correlation of Fixed Effects:
##         (Intr) relate pair
## relate  -0.138
## pair    -0.499  0.004
## readage -0.601  0.001  0.003

Anova(mask.model)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##           Chisq Df Pr(>Chisq)
## relate  16.2399  1  5.581e-05 ***
## pair     4.3233  1   0.037595 *
## readage  9.7997  1   0.001745 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

mask.model1=lmer(react ~ relate + pair + vocabgroup + (1 + pair|subject) + (1
+ relate|word), data=L2maskRT, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 1.390e-01  3.061e-04 453.963
## relate      3.393e-04  8.725e-05   3.888
## pair        1.105e-04  5.252e-05   2.104
## vocabgroup  9.212e-04  3.789e-04   2.431
##
## Correlation of Fixed Effects:
##            (Intr) relate pair
## relate     -0.097
## pair       -0.463  0.004
## vocabgroup -0.601  0.000 -0.003

Anova(mask.model1)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##              Chisq Df Pr(>Chisq)
## relate     15.1203  1  0.0001009 ***
## pair        4.4250  1  0.0354155 *
## vocabgroup  5.9117  1  0.0150409 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#check interactions
mask.IN=lmer(react ~ relate * pair + (1 + pair|subject) + (1|word),
data=L2maskRT, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 1.395e-01  2.826e-04 493.728
## relate      1.215e-04  2.230e-04   0.545
## pair        5.617e-05  7.404e-05   0.759
## relate:pair 1.089e-04  1.030e-04   1.057
##
## Correlation of Fixed Effects:
##             (Intr) relate pair
## relate      -0.399
## pair        -0.638  0.651
## relate:pair  0.369 -0.926 -0.702

Anova(mask.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                  Chisq Df Pr(>Chisq)
## (Intercept) 2.4377e+05  1     <2e-16 ***
## relate      2.9680e-01  1     0.5859
## pair        5.7550e-01  1     0.4481
## relate:pair 1.1176e+00  1     0.2904
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#test pair factor for unplanned comparisons
summary(m1<-aov(react~pair, data=L2maskRT))
##               Df   Sum Sq   Mean Sq F value Pr(>F)
## pair           2 0.000029 1.448e-05   2.688 0.0682 .
## Residuals   2142 0.011539 5.387e-06
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TukeyHSD(m1,"pair",ordered=TRUE)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = react ~ pair, data = L2maskRT)
##
## $pair
##             diff           lwr          upr     p adj
## 1-2 6.176838e-05 -0.0002268411 0.0003503779 0.8703790
## 3-2 2.709902e-04 -0.0000166189 0.0005585994 0.0697409
## 3-1 2.092218e-04 -0.0000782855 0.0004967292 0.2028556
```

```
#test simple effects in interaction
maskL2.1 <- rename(filter(L2maskRT,pair==1))
mask.simple=lmer(react ~ relate + (1 + relate|subject) + (1+relate|word),
data=maskL2.1, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 0.1395705  0.0002369 589.070
## relate      0.0003644  0.0002074   1.757
##
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.199
```

```
Anova(mask.simple)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 3.0874  1     0.0789 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
maskL2.2 <- rename(filter(L2maskRT,pair==2))
mask.simple1=lmer(react ~ relate + (1 + relate|subject) + (1 + relate|word),
data=maskL2.2, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 1.397e-01  2.765e-04 505.107
## relate      7.488e-05  1.744e-04   0.429
##
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.401
```

```
Anova(mask.simple1)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 0.1844  1     0.6676
```

```
maskL2.3 <- rename(filter(L2maskRT,pair==3))
mask.simple2=lmer(react ~ relate + (1 + relate|subject) + (1 + relate|word),
data=maskL2.3, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 0.1396810  0.0002175  642.31
## relate      0.0005828  0.0001911    3.05
```

```
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.245
```

```
Anova(mask.simple2)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 9.2996  1   0.002292 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Elementary Age Group

```
#test interaction model
mask.scoreIN <-glmer(score ~ relate * pair +  (1+relate|subject) + (1
+relate|word), data=Emask, family=binomial)
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.01143    0.31809   6.323 2.56e-10 ***
## relate       0.11500    0.41823   0.275    0.783
## pair         0.08732    0.13013   0.671    0.502
## relate:pair -0.05118    0.18597  -0.275    0.783
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##

## Correlation of Fixed Effects:
##             (Intr) relate pair
## relate      -0.672
## pair        -0.799  0.608
## relate:pair  0.559 -0.875 -0.700
```

```
#enter the model
mask.model=lmer(react ~ relate + pair + (1 + pair|subject) + (1|word),
data=EmaskRT, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 1.377e-01  3.618e-04 380.546
## relate      4.205e-04  1.180e-04   3.562
## pair        1.856e-04  7.536e-05   2.463
##
## Correlation of Fixed Effects:
##        (Intr) relate
## relate -0.165
## pair   -0.633  0.003
```

```
Anova(mask.model)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##          Chisq Df Pr(>Chisq)
## relate 12.6913  1  0.0003674 ***
## pair    6.0669  1  0.0137741 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#check interactions
mask.IN=lmer(react ~ relate * pair + (1 + pair|subject) + (1|word),
data=EmaskRT, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 1.378e-01  3.895e-04 353.657
## relate      2.489e-04  3.125e-04   0.796
## pair        1.427e-04  1.044e-04   1.367
## relate:pair 8.565e-05  1.443e-04   0.593
##
## Correlation of Fixed Effects:
##             (Intr) relate pair
## relate      -0.403
## pair        -0.681  0.642
## relate:pair  0.373 -0.926 -0.693
```

```
Anova(mask.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                 Chisq Df Pr(>Chisq)
## (Intercept) 1.2507e+05  1     <2e-16 ***
## relate      6.3430e-01  1     0.4258
## pair        1.8674e+00  1     0.1718
## relate:pair 3.5210e-01  1     0.5529
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#test pair factor for unplanned comparisons
summary(m1<-aov(react~pair, data=EmaskRT))
##               Df   Sum Sq   Mean Sq F value Pr(>F)
## pair           2 0.000031 1.542e-05   2.055  0.128
## Residuals   1539 0.011551 7.505e-06
TukeyHSD(m1,"pair",ordered=TRUE)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
```

```
## Fit: aov(formula = react ~ pair, data = EmaskRT)
##
## $pair
##              diff           lwr          upr     p adj
## 2-1 2.741780e-04 -1.279090e-04 0.0006762650 0.2460556
## 3-1 3.203430e-04 -7.960357e-05 0.0007202896 0.1450971
## 3-2 4.616502e-05 -3.545719e-04 0.0004469019 0.9605385
```

```
#test simple effects in interaction
maskE.1 <- rename(filter(EmaskRT,pair==1))
mask.simple=lmer(react ~ relate + (1 + relate|subject) + (1|word),
data=maskE.1, REML=FALSE)
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept) 0.1378318  0.0003632 379.478
## relate      0.0004255  0.0002176   1.955
##
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.472
```

```
Anova(mask.simple)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 3.8239  1    0.05053 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
maskE.2 <- rename(filter(EmaskRT,pair==2))
mask.simple1=lmer(react ~ relate + (1 + relate|subject) + (1|word),
data=maskE.2, REML=FALSE)
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept) 0.1382101  0.0002976 464.388
## relate      0.0002520  0.0002245   1.122
##
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.134
```

```
Anova(mask.simple1)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 1.2598  1     0.2617
```

```
maskE.3 <- rename(filter(EmaskRT,pair==3))
mask.simple2=lmer(react ~ relate + (1 + relate|subject) + (1|word),
data=maskE.3, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 0.1381024  0.0003086 447.513
## relate      0.0005859  0.0002062   2.841
##
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.489
```

```
Anova(mask.simple2)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 8.0725  1   0.004494 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

All Groups

```
#accuracy models
mask.scoreIN <-glmer(score~relate * pair + group + (1|subject) + (1|word),
data=allmask, family=binomial)

## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.673568   0.288685  16.189   <2e-16 ***
## relate       0.068687   0.271921   0.253    0.801
## pair         0.128375   0.090562   1.418    0.156
## group       -0.860928   0.096400  -8.931   <2e-16 ***
## relate:pair -0.002166   0.129797  -0.017    0.987
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) relate pair   group
## relate      -0.458
## pair        -0.601  0.643
## group       -0.671 -0.001 -0.004
## relate:pair  0.422 -0.921 -0.698  0.000
summary(m1<-aov(score~group, data=allmask))
##               Df Sum Sq Mean Sq F value Pr(>F)
## group          2   10.4   5.194     123 <2e-16 ***
## Residuals   8276  349.4   0.042
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 1 observation deleted due to missingness
```

```
TukeyHSD(m1,"group",ordered=TRUE)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = score ~ group, data = allmask)
##
## $group
##            diff          lwr        upr     p adj
## 2-3 0.081501437  0.066875980 0.09612689 0.0000000
## 1-3 0.090289414  0.076262564 0.10431626 0.0000000
## 1-2 0.008787977 -0.003210815 0.02078677 0.1988868
```

```
#enter the model
mask.IN=lmer(react ~ relate * pair * group + (1|subject) + (1|word),
data=allmaskRT, REML=FALSE)
## Fixed effects:
##                      Estimate Std. Error t value
## (Intercept)        1.414e-01  3.391e-04 417.078
## relate            -6.808e-04  3.041e-04  -2.239
## pair              -4.778e-05  9.939e-05  -0.481
## group             -1.102e-03  1.724e-04  -6.391
## relate:pair        5.165e-04  1.405e-04   3.675
## relate:group       3.065e-04  1.599e-04   1.918
## pair:group         5.315e-05  5.225e-05   1.017
## relate:pair:group -1.492e-04  7.385e-05  -2.020
## Correlation of Fixed Effects:
##             (Intr) relate pair   group  rlt:pr rlt:gr pr:grp
## relate      -0.449
## pair        -0.587  0.655
## group       -0.891  0.426  0.556
## relate:pair  0.415 -0.926 -0.707 -0.394
## relate:grop  0.412 -0.916 -0.600 -0.465  0.848
## pair:group   0.538 -0.600 -0.916 -0.607  0.648  0.656
## relt:pr:grp -0.380  0.848  0.648  0.430 -0.916 -0.926 -0.708
```

```
Anova(mask.IN, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
## Response: react
##                        Chisq Df Pr(>Chisq)
## (Intercept)       1.7395e+05  1  < 2.2e-16 ***
## relate            5.0112e+00  1  0.0251837 *
## pair              2.3110e-01  1  0.6307176
## group             4.0844e+01  1  1.649e-10 ***
## relate:pair       1.3509e+01  1  0.0002375 ***
## relate:group      3.6769e+00  1  0.0551712 .
## pair:group        1.0344e+00  1  0.3091175
## relate:pair:group 4.0821e+00  1  0.0433392 *

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(m1<-aov(react~group, data=allmaskRT))
##               Df  Sum Sq   Mean Sq F value Pr(>F)
## group          2 0.00460 0.0022979   403.3 <2e-16 ***
## Residuals   7872 0.04486 0.0000057
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
TukeyHSD(m1,"group",ordered=TRUE)
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = react ~ group, data = allmaskRT)
##
## $group
##             diff          lwr          upr p adj
## 2-3 0.0016951352 0.0015174365 0.0018728339 0e+00
## 1-3 0.0020364286 0.0018658023 0.0022070548 0e+00
## 1-2 0.0003412934 0.0001995935 0.0004829932 1e-07

#test simple effects for interactions
allmask4 <- rename(filter(allmaskRT,pair==1))
mask.simple3=lmer(react ~ relate + (1 + relate|subject) + (1 + relate|word),
data=allmask4, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 0.1395396  0.0001691 824.999
## relate      0.0002068  0.0001080   1.915
##

## Correlation of Fixed Effects:
##        (Intr)
## relate -0.045

Anova(mask.simple3)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 3.6675  1    0.05548 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
allmask5 <- rename(filter(allmaskRT,pair==2))
mask.simple4=lmer(react ~ relate + (1 + relate|subject) + (1 + relate|word),
data=allmask5, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 1.396e-01  1.588e-04 878.891
## relate      1.710e-04  9.223e-05   1.854
##
```

```
## Correlation of Fixed Effects:
##        (Intr)
## relate 0.185

Anova(mask.simple4)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 3.4383  1     0.0637 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
allmask6 <- rename(filter(allmaskRT,pair==3))
mask.simple5=lmer(react ~ relate + (1 + relate|subject) + (1 + relate|word),
data=allmask6, REML=FALSE)
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 0.1396361  0.0001508 925.972
## relate      0.0007070  0.0001105   6.397
## Correlation of Fixed Effects:
##        (Intr)
## relate -0.195

Anova(mask.simple5)
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: react
##         Chisq Df Pr(>Chisq)
## relate 40.925  1  1.582e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

allmask1 <- rename(filter(allmaskRT,group==1))
mask.simple=lmer(react ~ relate * pair + (1 + pair|subject) + (1|word),
data=allmask1, REML=FALSE)
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  1.401e-01  2.139e-04 655.106
## relate      -3.806e-04  1.729e-04  -2.202
## pair         1.562e-05  5.801e-05   0.269
## relate:pair  3.783e-04  7.987e-05   4.736
##
## Correlation of Fixed Effects:
##             (Intr) relate pair
## relate      -0.404
## pair        -0.533  0.637
## relate:pair  0.374 -0.926 -0.689
```

```
Anova(mask.simple, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                  Chisq Df Pr(>Chisq)
## (Intercept) 4.2916e+05  1  < 2.2e-16 ***
## relate      4.8482e+00  1    0.02767 *
## pair        7.2500e-02  1    0.78773
## relate:pair 2.2428e+01  1  2.181e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
allmask2 <- rename(filter(allmaskRT,group==2))
mask.simple1=lmer(react ~ relate * pair + (1 + pair|subject) + (1|word),
data=allmask2, REML=FALSE)
## Fixed effects:
##                Estimate Std. Error t value
## (Intercept)  1.397e-01  2.449e-04 570.523
## relate      -4.501e-05  1.987e-04  -0.227
## pair         3.699e-05  6.556e-05   0.564
## relate:pair  1.862e-04  9.178e-05   2.029
##
## Correlation of Fixed Effects:
##             (Intr) relate pair
## relate      -0.408
## pair        -0.622  0.652
## relate:pair  0.378 -0.926 -0.704
```

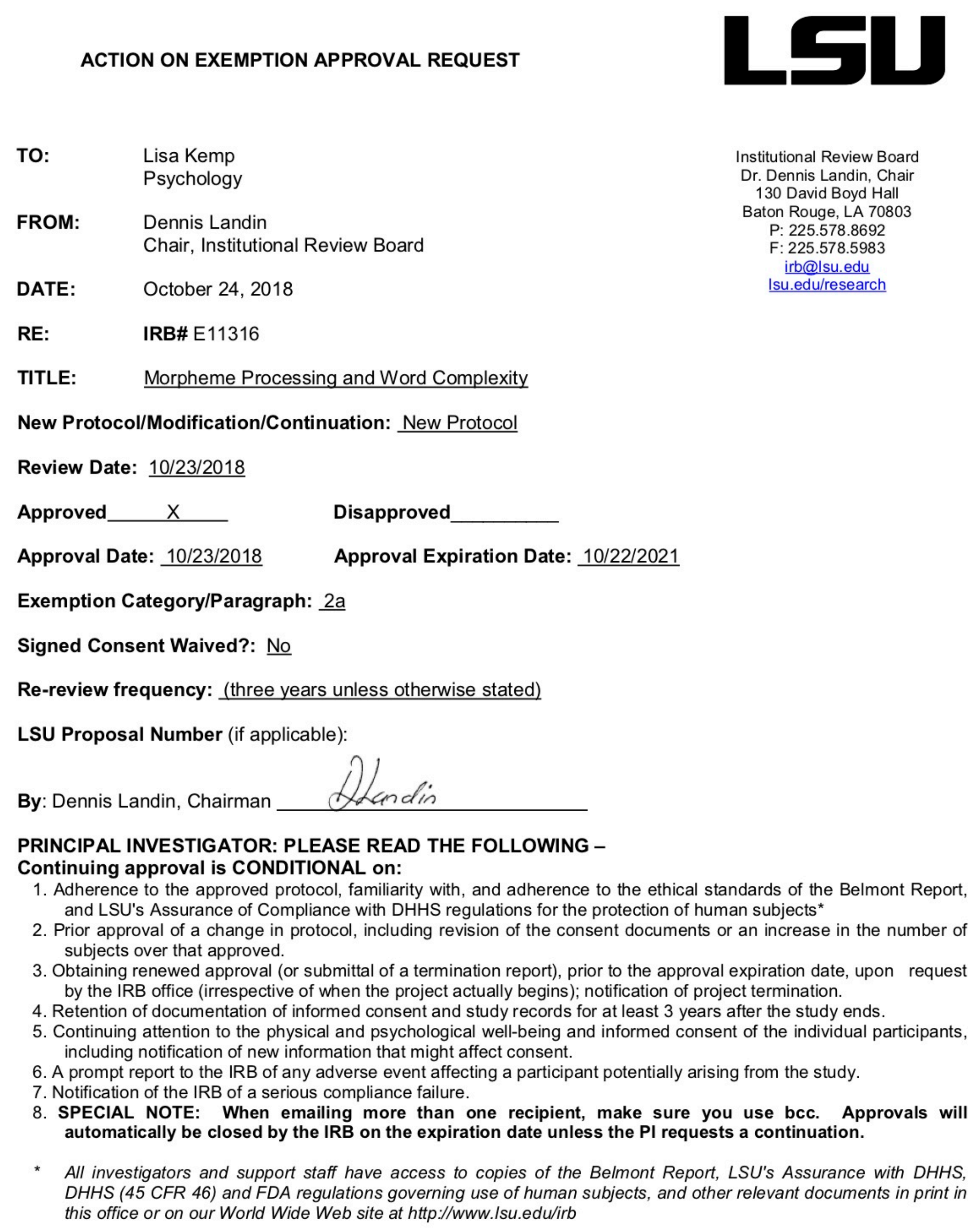
```
Anova(mask.simple1, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                  Chisq Df Pr(>Chisq)
## (Intercept) 3.2550e+05  1     < 2e-16 ***
## relate      5.1300e-02  1    0.82081
## pair        3.1840e-01  1    0.57256
## relate:pair 4.1168e+00  1    0.04246 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
allmask3 <- rename(filter(allmaskRT,group==3))
mask.simple2=lmer(react ~ relate * pair + (1 + pair|subject) + (1|word),
data=allmask3, REML=FALSE)
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept) 1.378e-01  3.895e-04 353.657
## relate      2.489e-04  3.125e-04   0.796
## pair        1.427e-04  1.044e-04   1.367
## relate:pair 8.565e-05  1.443e-04   0.593
##
```

```
## Correlation of Fixed Effects:
##             (Intr) relate pair
## relate      -0.403
## pair        -0.681  0.642
## relate:pair  0.373 -0.926 -0.693
```

```
Anova(mask.simple2, type=3)
## Analysis of Deviance Table (Type III Wald chisquare tests)
##
## Response: react
##                  Chisq Df Pr(>Chisq)
## (Intercept) 1.2507e+05  1     <2e-16 ***
## relate      6.3430e-01  1     0.4258
## pair        1.8674e+00  1     0.1718
## relate:pair 3.5210e-01  1     0.5529
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Appendix H. IRB Approval Forms



ACTION ON EXEMPTION APPROVAL REQUEST

LSU

Institutional Review Board
Dr. Dennis Landin, Chair
130 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu
lsu.edu/research

TO: Lisa Kemp
Psychology

FROM: Dennis Landin
Chair, Institutional Review Board

DATE: October 24, 2018

RE: **IRB#** E11316

TITLE: Morpheme Processing and Word Complexity

New Protocol/Modification/Continuation: New Protocol

Review Date: 10/23/2018

Approved X **Disapproved**__________

Approval Date: 10/23/2018 **Approval Expiration Date:** 10/22/2021

Exemption Category/Paragraph: 2a

Signed Consent Waived?: No

Re-review frequency: (three years unless otherwise stated)

LSU Proposal Number (if applicable):

By: Dennis Landin, Chairman DLandin

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. **SPECIAL NOTE: When emailing more than one recipient, make sure you use bcc. Approvals will automatically be closed by the IRB on the expiration date unless the PI requests a continuation.**

* *All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb*

ACTION ON EXEMPTION APPROVAL REQUEST

Institutional Review Board
Dr. Dennis Landin, Chair
130 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu
lsu.edu/research

TO: Lisa Kemp
Psychology

FROM: Dennis Landin
Chair, Institutional Review Board

DATE: January 22, 2019

RE: **IRB#** E11316

TITLE: Morpheme Processing and Word Complexity

New Protocol/Modification/Continuation: Modification

Brief Modification Description: Recruit adult non-native speakers of English, they will receive their own background questionnaire, added 10 new English vocabulary items to the vocabulary test.

Review date: 1/22/2019

Approved __X__ **Disapproved** __________

Approval Date: 1/22/2019 **Approval Expiration Date:** 10/22/2021

Re-review frequency: (three years unless otherwise stated)

LSU Proposal Number (if applicable):

By: Dennis Landin, Chairman DLandin

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. **SPECIAL NOTE: Make sure you use bcc when emailing more than one recipient. Approvals will automatically be closed by the IRB on the expiration date unless the PI requests a continuation.**

**All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb*

Institutional Review Board
Dr. Dennis Landin, Chair
130 David Boyd Hall
Baton Rouge, LA 70803
P: 225.578.8692
F: 225.578.5983
irb@lsu.edu
lsu.edu/research

ACTION ON PROTOCOL APPROVAL REQUEST

TO: Janet McDonald
Psychology

FROM: Dennis Landin
Chair, Institutional Review Board

DATE: February 18, 2019

RE: **IRB#** 4195

TITLE: Complex Word Processing in Elementary Age Children

New Protocol/Modification/Continuation: New Protocol

Review type: Full ___ Expedited _X_ **Review date:** 2/18/2019

Risk Factor: Minimal ___X___ Uncertain ______ Greater Than Minimal_______

Approved_____X_____ **Disapproved**__________

Approval Date: 2/18/2019 **Approval Expiration Date:** 2/17/2020

Re-review frequency: (annual unless otherwise stated)

Number of subjects approved: 60

LSU Proposal Number (if applicable)**:**

By: Dennis Landin, Chairman DLandin

PRINCIPAL INVESTIGATOR: PLEASE READ THE FOLLOWING –
Continuing approval is CONDITIONAL on:

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
7. Notification of the IRB of a serious compliance failure.
8. **SPECIAL NOTE: When emailing more than one recipient, make sure you use bcc.**

**All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at http://www.lsu.edu/irb*

References

Andrews, S., & Bond, R. (2009). Lexical expertise and reading skill: Bottom-up and top-down processing of lexical ambiguity. *Reading and Writing*, *22*(6), 687–711.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, *139*(2), 299–318.

Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language*, *68*(3), 279–296.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*.

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

Beyersmann, E., Casalis, S., Ziegler, J. C., & Grainger, J. (2015). Language proficiency and morpho-orthographic segmentation. *Psychonomic Bulletin & Review*, *22*(4), 1054–1061.

Beyersmann, E., Castles, A., & Coltheart, M. (2012). Morphological processing during visual word recognition in developing readers: Evidence from masked priming. *The Quarterly Journal of Experimental Psychology*, *65*(7), 1306-1326.

Beyersmann, E., & Grainger, J. (2018). Support from the morphological family when unembedding the stem. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(1), 135–142.

Beyersmann, E., Ziegler, J. C., Castles, A., Coltheart, M., Kezilas, Y., & Grainger, J. (2016). Morpho-orthographic segmentation without semantics. *Psychonomic Bulletin & Review*, *23*(2), 533–539.

Bracken, B. A., & Murray, A. M. (1984). Stability and predictive validity of the PPVT-R over an eleven month interval. *Educational & Psychological Research*.

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411.

Brooks, P. J., Kwoka, N., & Kempe, V. (2017). Distributional Effects and Individual Differences in L2 Morphology Learning: Determinants of L2 Morphology Learning. *Language Learning*, *67*(1), 171–207.

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

Carlisle, J. F. (2007). Fostering morphological processing, vocabulary development, and reading comprehension. *Vocabulary acquisition: Implications for reading comprehension*, 78-103.

Casalis, S., Commissaire, E., & Duncan, L. G. (2015a). Sensitivity to morpheme units in English as L2 word recognition. *Writing Systems Research*, *7*(2), 186–201.

Casalis, S., Quémart, P., & Duncan, L. G. (2015b). How language affects children's use of derivational morphology in visual word and pseudoword processing: evidence from a cross-language study. *Frontiers in Psychology*, *6*.

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, *97*(3), 165–182.

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*(1), 21–43.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*(4), 335–359.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, 33*, 497-505.

Coughlin, C. E., & Tremblay, A. (2015). Morphological decomposition in native and non-native French speakers. *Bilingualism: Language and Cognition*, *18*(03), 524–542.

Dawson, N., Rastle, K., & Ricketts, J. (2018). Morphological effects in visual word recognition: Children, adolescents, and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 645–654.

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

Diependaele, K., Sandra, D., & Grainger, J. (2005). Masked cross-modal morphological priming: Unravelling morpho-orthographic and morpho-semantic influences in early word recognition. *Language and Cognitive Processes*, *20*(1–2), 75–114.

Duncan, L. G., Gray, E., Quémart, P., & Casalis, S. (2011). Do good and poor readers make use of morphemic structure in English word recognition? *Journal of Portuguese Linguistics*, *10*(1), 143.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.

Feldman, L. B., Milin, P., Cho, K. W., Moscoso del Prado Martín, F., & O'Connor, P. A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience*, *9*.

Fowler, C. A., Napps, S. E., & Feldman, L. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory & Cognition*, *13*(3), 241–255.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(04).

Giraudo, H., & Dal Maso, S. (2016). The salience of complex words and their parts: Which comes first? *Frontiers in psychology*, *7*, 1778.

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.

Grainger, J., & Beyersmann, E. (2017). Edge-Aligned Embedded Word Activation Initiates Morpho-orthographic Segmentation. In *Psychology of Learning and Motivation* (Vol. 67, pp. 285–317).

Haxby, J. V., Parasuraman, R., Lalonde, F., & Abboud, H. (1993). SuperLab: General-purpose Macintosh software for human experimental psychology and psychological testing. *Behavior Research Methods, Instruments & Computers, 25*(3), 400-405.

Heyer, V. & 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*(03), 543–550.

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

Jared, D., Jouravlev, O., & Joanisse, M. F. (2017). The effect of semantic transparency on the processing of morphologically derived words: Evidence from decision latencies and event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(3), 422–450.

Lavric, A., Clapp, A., & Rastle, K. (2007). ERP evidence of morphological analysis from orthography: A masked priming study. *Journal of Cognitive Neuroscience*, *19*(5), 866-877.

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 12–31.

Lewis, G., Solomyak, O., & Marantz, A. (2011). The neural basis of obligatory decomposition of suffixed words. *Brain and Language*, *118*(3), 118–127.

Liang, L., & Chen, B. (2014). Processing morphologically complex words in second-language learners: The effect of proficiency. *Acta Psychologica*, *150*, 69–79.

Li, J., Taft, M., & Xu, J. (2017). The Processing of English Derived Words by Chinese-English Bilinguals: Chinese-English Bilinguals. *Language Learning*, *67*(4), 858–884.

Li, P., Zhang, F., Tsai, E. & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition,* 17(3), 673-6 80.

Liang, L., & Chen, B. (2014). Processing morphologically complex words in second-language learners: The effect of proficiency. *Acta Psychologica*, *150*, 69–79.

Longtin, C.-M., & Meunier, F. (2005). Morphological decomposition in early visual word processing☆. *Journal of Memory and Language*, *53*(1), 26–41.

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

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.

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

Masterson, J., Stuart, M., Dixon, M., & Lovejoy, S. (2010). Children's printed word database: Continuities and changes over time in children's early reading vocabulary. *British Journal of Psychology*, *101*(2), 221-242.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

Medeiros, J., & Duñabeitia, J. A. (2016). Not everybody sees the ness in the darkness: Individual differences in masked suffix priming. *Frontiers in Psychology*, *7*, 1585.

Napps, S. E. (1989). Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory & Cognition*, *17*(6), 729–739.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, *47*(4), 1343–1355.

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Retrieved from http://www.pstnet.com.

R Core Team (2017). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL https://www.R-project.org/.

Rastle, K. (2016). Visual word recognition. In *Neurobiology of Language* (pp. 255-264). Academic Press.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, *23*(7–8), 942–971.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*(6), 1090–1098.

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), 1793–1820.

Schmidtke, D., Van Dyke, J. A., & Kuperman, V. (2018). Individual variability in the semantic processing of English compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(3), 421–439.

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, *11*(02).

Solomyak, O., & Marantz, A. (2010). Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience*, *22*(9), 2042–2057.

Tamminen, J., Davis, M. H., & Rastle, K. (2015). From specific examples to general knowledge in language learning. *Cognitive Psychology*, *79*, 1–39.

Tsang, Y.-K., & Chen, H.-C. (2014). Activation of morphemic meanings in processing opaque words. *Psychonomic Bulletin & Review*, *21*(5), 1281–1286.

Tsang, Y.-K., Wong, A. W.-K., Huang, J., & Chen, H.-C. (2014). Morpho-orthographic and morpho-semantic processing in word recognition and production: Evidence from ambiguous morphemes. *Language, Cognition and Neuroscience*, *29*(5), 543–560.

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of memory and language*, *28*(6), 649-667.

Voga, M., Anastassiadis-Symeonidis, A., & Giraudo, H. (2014). Does morphology play a role in L2 processing? Two masked priming experiments with Greek speakers of ESL. Lingvisticae Investigationes, 37 (2).

Vita

Lisa Kemp began her career in Cognitive Psychology in 2011 when she enrolled in a Linguistics Masters program at California State University, Long Beach. This decision was driven by a curiosity in the functions and structure of human language. Lisa decided to pursue a Ph.D. in Psychology at Louisiana State University after her Masters thesis which formed a passion for experimental design involving questions related to language learning and use. She plans to pursue a career in data science following graduation.