2017-01-01

Bilingual Novel Word Learning In Sentence Contexts

Justin G. Lauro
University of Texas at El Paso, jglauro@utep.edu

BILINGUAL NOVEL WORD LEARNING IN SENTENCE CONTEXTS

JUSTIN GERALD LAURO

Doctoral Program in Psychology

APPROVED:

Ana Schwartz, Ph.D., Chair

Ashley Bangert, Ph.D.

Carla Contemori, Ph.D.

Wendy Francis, Ph.D.

Michael Zarate, Ph.D.

Charles Amber, Ph.D.
Dean of the Graduate School

BILINGUAL NOVEL WORD LEARNING IN SENTENCE CONTEXTS

by

JUSTIN GERALD LAURO, M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

THE UNIVERSITY OF TEXAS AT EL PASO

August 2017

Abstract

According to the *Instance-based framework* of adult vocabulary learning, each time a learner encounters a new word in context, an episodic memory trace, containing the word itself and the surrounding context, is formed. Previous research has demonstrated an advantage in memory for words encountered in a variety of semantic contexts. This advantage occurs because distinctive contexts generate more distinctive memory traces. While research has demonstrated that a variety of contextual characteristics are encoded (e.g. semantic context, modality, and language), it is unknown whether varying the language context in which a word is encountered has a similar benefit on memory for new words. Across two experiments, highly proficient Spanish-English bilinguals studied rare cognates in sentences in both languages. At test, they were asked to generate meanings and recognize semantically related and unrelated word pairs containing studied words. In Experiment 1, participants studied words and were tested visually. In Experiment 2, participants studied words and were tested auditorily. Across both experiments, an advantage in generating meanings of words learned in different semantic contexts was observed in both the L1 and L2. However, no additive effects of language variation of encounters with words at study were directly observed. This is the first study to explicitly test the context variability hypothesis in bilinguals' L2, as well as in the auditory modality. The results are discussed within an *instance-based* theoretical perspective, drawing from literature on bilingual conceptual access.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

There is a growing body of research examining factors that foster the acquisition of new words. One factor is the variety of contexts in which a novel word is encountered (Bolger, Balass, Landen & Perfetti, 2008; Balass, 2011; Elgort & Piasecki, 2014; Elgort, Perfetti, Rickles, & Stafura, 2015). According to the *Instance-based framework* of adult vocabulary learning (Reichle & Perfetti, 2003; Bolger et al., 2008), each time a learner encounters a new word in context, a memory trace, which includes the word itself and the surrounding context, is formed. Encountering words in varied contexts is advantageous because distinctive contexts generate more distinctive memory traces. While research demonstrates that a variety of contextual characteristics are encoded (e.g. semantic context, modality), there are currently no published studies that have examined whether variety in the language of context in which a word is encountered has a similar benefit on memory for new words. This is a critical gap in the literature since bilinguals encounter novel words in two languages. The goal of the present study was to examine whether language membership of context acts as a distinctive cue in the episodic memory trace.

1.1 Instance-Based Framework of Adult Vocabulary Learning

A lexicon is made up of a network of connections between word forms ("lexical representations") and meanings ("semantic representations"). A new word becomes part of the lexicon when its lexical-semantic representation is strongly connected to the semantic representations and features of other related, known words within the lexicon's network (Bolger et al., 2008; Masson, 1995; McRae, de Sa, & Seidenberg, 1997; Plaut & Booth, 2000). Each exposure to a word strengthens the lexical representation (Morton, 1970; McClelland & Rumelhart, 1981) or the association between the lexical and semantic representation. In this

way, word learning is incremental, increasing with each encounter of a word in a given context (Bolger et al., 2008; Fukkink, Blok, & de Glopper, 2001; Jenkins, Stein, & Wysocki, 1984; Nagy, Anderson, & Herman, 1987). Over a lifetime of language use, these word learning experiences produce a large number of memory traces, and a large vocabulary (Reichle & Perfetti, 2003).

The *Instance Based Framework* (Reichle & Perfetti, 2003) provides an account of how these representations are formed through encounters with words. Its core principles are derived from more general *instance-based* and *exemplar* models (Logan, 1988; Hintzman, 1984; Medin & Schaffer, 1978; Smith & Zarate, 1992) which assume that each encounter with a stimulus is encoded, stored and retrieved as a separate memory trace (Logan, 1988). These traces are then activated and potentially retrieved during later encounters with the word. Episodic information, such as modality (e.g. Nelson, Balass, & Perfetti, 2005) is also encoded in these traces. For example, learners are more accurate and more confident in recognizing words when they are tested in the same modality as they are studied. Even minute perceptual details such as the typeface in which a word was read (Marsolek, Schacter, & Nicholas, 1996; Marsolek, 2004) or its location on a screen are stored in memory (Dufau, Grainger, & Holcomb, 2008). When novel words are encountered, orthographic and phonological features of the word are encoded and bound together in the memory trace with semantic information that is specific to that particular context. (Bolger et al., 2008; Reichle & Perfetti, 2003). For example, when learning the word *tiger*, a reader may learn that a tiger "has fur", and that other animals "have fur", and these animals also tend to have claws and a tail and be carnivorous (McRae et al., 1997).

There are two mechanisms, an active mechanism and a resonance mechanism, which can act independently or in unison to encode and retrieve those traces. The active mechanism enables

learners to attend, search, retrieve and evaluate information about a word meaning and is the basis of intentional word learning (McKeown, 1985). These cognitive operations consume significant working memory resources (Daneman & Green, 1986), and less-skilled comprehenders struggle to learn word meanings from context (McKeown, 1985) even after several exposures (Jenkins et al., 1984; van Daalen-Kapteijns et al., 2001).

Although intentional learning results in superior memory for words, the majority of new words are learned incidentally (Nagy & Anderson, 1984), which is based on the resonance mechanism. These resonance processes consist of the automatic activation of previously encoded episodic traces in memory in which the word was encountered in context, as well as other existing words that are related to the context. Through this resonance mechanism readers and listeners can automatically determine the meaning of the previously encountered novel words. Both of these mechanisms can function simultaneously. Learners can consciously attempt to extract the meaning of an unknown word from the given context, while the resonance mechanism retrieves previous memory traces of encounters with that word. If readers have encountered a novel word in various contexts previously, the resonance mechanism has more memory traces that may potentially become activated.

The *Context Variability Hypothesis* (Balass, 2011; Bolger et al., 2008) posits that encountering words in a variety of contexts leads to better learning of the meanings of words than when words are encountered repeatedly in the same context. This advantage occurs because each experience with a word creates a unique episodic memory trace that includes context-specific information, such as the sentence context, as well as context-independent information, such as the semantic meaning of the target word (Reichle & Perfetti, 2003; Nelson et al., 2005; Perfetti, Wlotko, & Hart, 2005). Unique experiences with a particular word establish a network

of connections around a lexical-semantic representation. Thus, when encountering a recently learned word in a new context, the resonance mechanism activates previously encoded, related memory traces, while the new exposure to the word is also encoded in that representation network. Repeated exposure to a word in the same context strengthens that one specific memory trace. If subsequent experiences with a word do not vary in context, the original context in which the word was encountered remains a strong component of the memory trace. Thus, words may be more recognizable when encountered repeatedly in the same specific context. However, encountering a word in a variety of contexts decreases the dependency on context-specific information by providing more features of the word in memory, enabling the formation of a more robust, "decontextualized" representation (Balass, 2011; Nelson, et al., 2005).

The effects of context variability on word learning have been observed in both incidental (Elgort et al., 2015) and intentional learning (Balass, 2011; Bolger et al., 2008). For example, in a study investigating the effects of context on word learning, learners were exposed to rare words in either four-varying sentences, or with a dictionary definition (Balass, 2011). After the initial learning phase, participants were exposed to new sentences in which the final word was one of the rare-learned words, or a non-studied control word. Learners were more accurate in deciding whether the target word fit in that context when the target word was originally presented in four-varied sentences. Learners were also able to generate more definitions of recently learned words when they were studied in varying contexts (Balass, 2011).

The advantage of studying novel words in varying sentence contexts has been observed even when the words' definitions are provided in addition to sentences (Bolger et al., 2008). However, the benefit of context variability was attenuated when definitions were provided. The interpretation was that, when a definition is provided in addition to the sentence context, each

encounter is essentially less distinct from each other, and therefore the effects of semantic variability of the provided sentence context were attenuated, but not absent.

Another key finding from the literature is that context variability effects have been observed even in low-constraint sentence contexts, when the meaning of the target word was not explicitly stated in the sentences. Variability of semantic context facilitated semantic congruency decisions even in less semantically “helpful” contexts (Balass, 2011). This furthers the idea that variability of each encounter, not only in semantically rich contexts, can facilitate knowledge acquisition. These findings demonstrate the importance of contextual variability in the acquisition of vocabulary knowledge, as well as the ability to use and recognize new vocabulary in novel situations or environments. Encountering a word in a variety of contexts appears to create a robust lexical-semantic network of memory traces which are readily available for the resonance mechanism to retrieve when encountering that word again.

1.1.1 Lexical Quality Hypothesis

Lexical properties of the word itself also influence the strength of memory traces for recently encountered words. The *lexical quality hypothesis* posits that the strength of connections between word forms and meanings affects retrieval of those word representations from memory (Perfetti, 1985, 1992; Perfetti & Hart, 2001, 2002). High quality lexical representations have strong connections from form to meaning, and are tightly bound together (Perfetti, 2007). Multiple encounters with a word produce a single core representation, consisting of orthographic, phonological, and semantic information components. Each of these components is subject to constraint from its relative system (e.g. phonology, orthography, and syntax/semantics). A word that does not specify the value of one of its components will be of lower quality (Perfett & Hart, 2001). For example, a homonym such as *record*, will have

ambiguity at the phonological level, and thus a lower lexical quality. A student may have a *record* of their academic transcripts and he might *record* a lecture from his introduction to psychology class. A word is of high lexical quality if it contains both semantic and phonological information to locate that specific memory trace (i.e. only the intended representation), and consistently locates the trace (Perfetti, 1985). In other words, a lexical representation has high quality to the extent that it has a fully specified orthographic representation, and redundant phonological representations (from spoken language and from orthographic-to-phonological mappings). High lexical quality facilitates retrieval of the intended representation from memory (Perfetti, 1985; 1992).

Lexical quality is also influenced by lexical proficiency of the reader. Skilled readers have a greater probability of adding new information (about spelling, pronunciation, or meaning) to a low-quality representation. For example, many low frequency or specialized vocabularies will have low lexical quality for both skilled and less-skilled readers. However, skilled readers have an advantage in cognitive resources (such as decoding skills, and spelling and grammar skills). Because of these advantages in comprehension resources, skilled-readers can maximize impoverished experiences with new representations and learn from novel exposures more-so than less-skilled readers. Skilled-readers thus have many high quality word representations, while less-skilled readers have fewer high quality word representations (Hart & Perfetti, 2008).

If lexical quality is acquired through experiences with words, the acquisition of lexical quality can be observed in novel word learning by comparing less-skilled and more-skilled readers. Differences in time spent reading, or number of encounters with a particular word, predict the degree of lexical quality differences between high and less-skilled readers (Hart & Perfetti, 2008). Therefore, a high frequency word (according to a corpus database) may actually

be experienced less frequently for less-skilled readers compared to more skilled readers. A low frequency word may be functionally even lower frequency for a less-skilled reader (Hart & Perfetti, 2008; Roth & Perfetti, 1980). Because of the subjective frequency of words, highly skilled readers are better able to recognize recently learned words compared to less skilled readers (Perfetti et al., 2005).

1.1.2 Second Language Vocabulary Learning

The logic behind the *Instance-Based Framework* has been extended to the learning of words in a second language (L2). Bilinguals or L2 learners can learn the meanings of novel L2 words intentionally (Elgort, 2011; Elgort & Piasecki, 2014; van Hell & Mahn, 1997), or incidentally in context (Elgort et al., 2015). For example, in an intentional L2 word learning study, bilinguals studied L2 pseudowords throughout a one week period. At test, participants demonstrated both acquired lexical and semantic knowledge of the recently learned pseudowords and were able to integrate both lexical and semantic information about the newly learned L2 pseudowords into their existing lexical-semantic networks (Elgort, 2011).

A seminal study on L2 incidental word learning in context demonstrated how bilinguals acquire word knowledge from context, and how the presence of context can activate connections to already known words (Elgort et al., 2015). Highly proficient bilinguals were able to correctly identify semantically related word pairs faster than semantically unrelated word pairs when the critical word was originally presented in distinct sentences at the study phase and embedded in a new sentence during the testing phase a day later. This finding suggests that encountering a rare or novel L2 word in a new context activates L2 access to the initially established semantic features of the previously learned word, along with the possible addition of new features (Elgort et al., 2015). Encountering a novel L2 word in a sentence context stimulated memory retrieval

of the previous encounters with that word and its associated semantic features. Although these findings inform and extend the *Context Variability Hypothesis* (Bolger et al., 2008) to L2 word learning, it was not the main goal of that study to explicitly test this hypothesis. During the study phase, all words were presented in 3 distinct semantic contexts. As such, it remains unknown whether facilitated access to related word-pairs was due directly to the increased activation of connections to known words associated with encountering a word across multiple, distinct semantic contexts.

Furthermore, a unique characteristic of word learning for bilinguals is that they can make explicit connections between L2 form and L1 translation equivalents to aid in L2 vocabulary learning. For example, one study found that highly proficient bilinguals established lexical and semantic representations of novel pseudowords when those pseudowords were originally presented with their L1 "translation equivalents" (Elgort & Piasecki, 2014). The strength of episodic traces of L2 word encounters is affected by learners' L2 lexical semantic knowledge (Elgort et al., 2015) and the ability to learn L2 words incidentally is modulated by proficiency in the L2 (Elgort et al., 2015; Elgort & Warren, 2014; Ferrell-Tekmen & Daloğlu, 2008; Horst, Cobb, & Meara, 1998; Pulido & Hambrick, 2008). Less proficient bilinguals or L2 learners with limited L2 vocabulary, fewer lexical-semantic connections to other known words may be limited (Elgort et al., 2015; Rodriguez-Fornells, Cunillera, Mestres-Misse, & de Diego-Balaguer, 2009).

If the process of learning the meanings of new words involves strengthening the links between semantic features within a network, a newly acquired representation will be integrated into an existing lexical-semantic memory network. Highly proficient bilinguals were able to use previously established lexical-semantic connections in addition to the contextual information in which a L2 novel word is encountered. Less proficient bilinguals may have insufficiently

developed L2 lexical-semantic networks from which to make connections to the meanings of novel words. Furthermore, less proficient bilinguals may lack the automaticity to access L2 lexical-semantic representations. Therefore, the episodic traces for each encounter with a novel word in context are stronger when L2 lexical semantic knowledge is greater.

Parallels to highly proficient and less-proficient readers' lexical quality for learned words may also be drawn. Although the lexical quality hypothesis does not explicitly mention bilinguals, form to meaning mappings may be between- or within-languages. If the lexical quality of words depends on tightly bound form to meaning mappings (Perfetti, 2007), L2 novel words will have fewer connections to their L1 translation equivalents, and thus be of lower lexical quality. Additionally, the lexical quality of L2 words may be lower for less proficient bilinguals, who may have fewer encounters with L2 words in general. Therefore, they may have applied more cognitive resources to lower level comprehension skills (e.g. decoding skills or grammar skills), limiting the ability to extract meaning from context.

1.2 Language as a Context in Novel Word Learning

Languages differ along a variety of dimensions. Even languages that share the same orthographic script (e.g. English and Spanish), have different phonological and syntactic rule systems. The first study to directly investigate whether the language membership of words is integrated into episodic memory compared the recall of target words that were studied across within-language or between-language repetitions (Glanzer & Duarte, 1971). Individual target words were studied across multiple trials in either the same or different languages. The study found that encountering a word once in each of two different languages led to better recall performance than encountering a word twice in a single language, but only when the lag between repetitions during the study phase was short; as the lag between repetitions increased, within-

and between-language repetitions were equally as effective. This pattern suggests that the salience of language form features versus semantic features differs according to the strength and duration of the memory trace (Francis, 1999). When the repetition of target words occurs at shorter lags, non-semantic, information within the trace may be more salient, whereas non-semantic information may be forgotten or less active at longer repetition intervals.

Even though memory for the language of input is not a necessary component of encoding conceptual information, bilinguals are typically good at memory for the language in which they encountered a concept, even when the learning is incidental (Strobach, 2015). Equivalent memory for language of input has been observed when words are studied in isolation or in sentence contexts (Batchelder & Riefer, 1990; Strobach, 2015), as well as following visual or auditory encoding (Strobach, 2015). These findings suggest that the language of the surrounding context is indeed encoded in episodic traces during word encounters. However, to what extent distinct language contexts have on learning words meanings is still unknown.

1.2.1 Visual or Auditory Presentation

If episodic traces of word learning events include context-specific information, it is likely that modal information (i.e. written or spoken) is also included in the trace. There has been some evidence to suggest that visual stimuli have an advantage in memory compared to auditory stimuli (Dean, Yekovich, & Gray, 1988; Gallo, McDermott, Percer & Roediger, 2001; Nelson et al., 2005). This may be due, at least in part, to the activation of phonological information during the visual presentation of words (Lukatela & Turvy, 1994; Lukatela, Frost, & Turvey, 1999; Naish, 1980; Perfetti, Bell, & Delaney, 1988; Perfetti & Bell, 1991; Unsworth & Pexman, 2003; Van Orden, Johnston, & Hale, 1988). For example, imagine someone reading silently (or overtly) and mouthing the words they are reading. This cross-modal code activation has also

been observed for auditory input (Castles, Holmes, Neath, & Kinoshita, 2003; Seidenberg & Tanenhaus, 1979; Slowiaczek, Soltano, Weiting, & Bishop, 2003; Ziegler & Ferrand, 1998). However, orthographic-to-phonological activation is stronger than phonological-to-orthographic activation (Borowsky, Owen, & Fonos, 1999). Thus it is more likely that a memory trace for a visually learned word may contain both orthographic and phonological information than a word learned auditorily.

Indeed, a study comparing the modality of input during a training phase of rare English words found that orthographic representations were learned faster, and recognized more accurately than phonological representations (Nelson et al., 2005). Although previous research has directly compared visual and auditory presentation of rare words in single language word learning studies, no one has examined the influence of modality on word learning in cross-linguistic contexts in bilinguals. Bilinguals are good at recalling in which language they have encountered a word in both visual and auditory modalities (Strobach, 2015; Rose, Rose, King, & Perez, 1975). Therefore, language membership information should influence retrieval of context-specific information when words are encountered visually and auditorily.

1.3 Bilingual Conceptual Memory

When postulating the processes of encoding new word meanings, it is critical to consider the unique characteristics of the bilingual lexical-conceptual system. A key difference between monolinguals and bilinguals is that the latter have at least two labels for a single concept. It has been well established in the literature that translation equivalents across languages are linked to a single, shared conceptual store (for reviews, see Francis, 1999; Kroll & Tokowicz, 2005). Evidence for a shared semantic system comes from cross-language priming studies, which have observed both within- and between-language repetition priming effects (de la Riva, Francis, &

Garcia, 2012; Francis, Fernandez & Bjork, 2010; Francis & Goldmann, 2011; Taylor & Francis, 2017).

The *Distributed feature model* (*DFM*) (de Groot, 1992) is of particular relevance to the present study because it focuses on specific connections between lexical form features and semantic representations. According to this model, words are made up of a number of lexical and semantic features. These features are represented as nodes in an interactive network (see Figure 1.1). One assumption of this model is that the degree of semantic feature overlap across languages affects access to the semantic representation. Some features may have strong links to both languages, whereas other features may have weaker connections to one of the languages. The degree of semantic feature overlap across languages is determined by linguistic properties of the word. For example, concrete words and cognates have a greater number of features that are strongly linked to both languages relative to abstract words and non-cognates (van Hell and de Groot, 1998). The specific combination of features that are activated during semantic access can vary between languages. Words that share more features, lexical and semantic, across languages, such as cognates and concrete nouns, are accessed more quickly than words that do not share features across languages (e.g. noncognates and abstract words) (van Hell and de Groot, 1998).

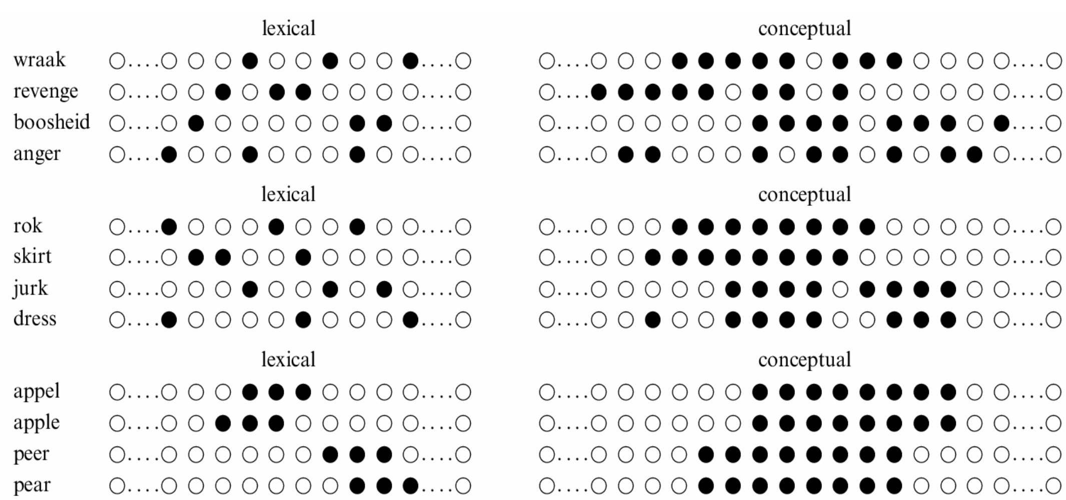

Figure 1. Distributed representation model (Reprinted from van Hell, 1998).

The assumptions of the *DFM* may be applied to novel word learning. If a word representation is made up of connections from lexical-to-semantic features of a word, words that are learned in two languages may have more semantic features shared across languages than a word learned in only one language. If the novel words are also cognates, we would expect a higher degree of shared lexical features as well. Therefore, bilinguals should have facilitated access to the meanings of recently learned words that were studied in both of their languages.

1.4 The Present Study

The present study investigated the influence of two types of contextual variation; 1) language of the context, and 2) the meaning conveyed by the context, on the subsequent memory of novel words. Across two experiments, participants studied lists of rare English-Spanish cognates in either the same repeated sentence context, or across varying semantic contexts. At test, they were asked to make semantic relatedness judgments on the studied (and non-studied) items, as well as generate meanings of those items. The central hypothesis of the present study is that encountering words in different languages make the episodic traces more distinctive, which will in turn result in a more robust lexical-semantic network of connections surrounding each word representation. Therefore, we predict that bilinguals' ability to learn the meanings of novel words should be improved when words are presented in two languages relative to one language. A second hypothesis is that distinctiveness of language contexts will vary depending whether words are encountered visually or auditorily. Language-specific phonological cues might provide an extra layer of distinctiveness for auditory word learning, whereas the language pair used in the present study, English and Spanish, are very similarly orthographically, and therefore language encounters across languages might not be as distinct. We also predicted that the

semantic variation effect that has previously been observed in single language word learning studies (e.g. Bolger et al., 2008, Balass, 2011) would be replicated in a bilingual sample. Although previous research on L2 vocabulary learning has examined the effects of varying semantic contexts (e.g. Elgort et al., 2015) as evidence for the formation of lexical-semantic network, to our knowledge no L2 vocabulary studies have directly compared word learning in varying semantic contexts and repeated semantic context.

Experiment 1 tested these hypotheses when words were presented visually. Experiment 2 tested these hypotheses when words were presented auditorily. A cross-experimental comparison between Experiments 1 and 2 will allow a direct comparison between visual and auditory stimuli encountered in the same semantic context conditions, and language context conditions. We expected an overall advantage in visual learning of words compared to auditory word learning. However, we expected the effect of language variation to be more pronounced when stimuli were presented auditorily than when they were presented visually due to the greater phonological distinctiveness across English and Spanish relative to orthographic distinctiveness across these languages.

Chapter 2: Experiment 1

Experiment 1 tested the hypothesis that variation in the language of context in which novel words are studied will increase the distinctiveness of the episodic memory traces formed from those encounters. Therefore, novel word learning should be improved when words are studied in two languages compared to only one language. A secondary goal of this experiment was to replicate the advantage of varied semantic contexts observed for monolinguals monolinguals (Bolger et al., 2008; Balass, 2011). Although context variability advantages have been studied in L2 speakers, previous studies did not include a condition in which the new words are presented in the same, repeated context in an L2 (Elgort et al., 2015).

2.1 Method

2.1.1 Participants

Seventy-two highly proficient Spanish-English bilingual speaking undergraduates (56 females, 16 males) from the University of Texas at El Paso participated for course credit. The median age of participants was 19. All participants learned Spanish first. The mean AoA for English was 6.19 (range 1-16), while the mean AoA for Spanish was 1.35 (range birth-5).

Language Proficiency. Participants' language proficiency was assessed using a subjective measure (ESPADA: Francis & Strobach, 2013) and objective measure (WMLS-R: Woodcock-Muñoz: Woodcock et al., 2005) (see materials section for a description of these instruments). Overall, the mean age equivalency scores in English and Spanish was not statistically different for picture vocabulary (English M = 12.20; Spanish M = 11.20) and reading comprehension (English M = 15.17; Spanish M = 15.10). Participants were classified as English-dominant or Spanish-dominant based on their relative Picture Vocabulary scores (see Table 2.1). Whichever language participants scored higher on was classified as their dominant language.

Based on this criterion, 40 participants were classified as English dominant, and 32 were Spanish dominant. The two language dominance groups did not differ significantly in terms of L1 passage comprehension, $t(70) = 1.88$, $p = .06$, or L1 picture vocabulary, $t(70) = -.186$, $p = .85$; nor did the language dominance groups differ in terms of L2 passage comprehension, $t(70) = .542$, $p = .59$, or L2 picture vocabulary, $t(70) = .005$, $p = .99$.

Self-ratings on the subjective measure aligned with the objective classification of English versus Spanish dominant. English dominant participants rated themselves as having better proficiency in English than in Spanish, whereas Spanish-dominant participants rated themselves as having better proficiency in Spanish than in English.

Table 2.1: Participant Language Characteristics

	Experiment 1		**Experiment 2**	
Characteristic	**English Dominant**	**Spanish Dominant**	**English Dominant**	**Spanish Dominant**
N	40	32	39	33
Age (Median)	19	19	19	20
AoA English	5.01	7.69	4.36	9.45
AoA Spanish	1.45	1.22	2.23	1.27
Objective Ratings				
AE English Picture Vocabulary	14.00	9.95	15.38	9.87
AE Spanish Picture Vocabulary	9.95	12.76	9.86	13.00
AE English Passage Comprehension	17.42	12.36	18.24	13.31
AE Spanish Passage Comprehension	13.00	17.72	10.36	16.39
Subjective Ratings (1-10)				
English Reading	8.77	8.16	8.95	7.55
English Writing	8.72	7.99	8.72	7.52
English Speaking	8.95	7.96	8.95	7.85

English Listening	9.18	9.03	9.59	8.67
Spanish Reading	7.33	8.35	7.05	8.88
Spanish Writing	6.28	7.58	6.15	8.30
Spanish Speaking	8.39	9.21	7.51	9.39
Spanish Listening	8.71	9.47	8.56	9.24

2.1.2 Materials

Critical word stimuli. An original pool of 212 English-Spanish cognates with English CELEX frequencies less than 7.5 words per million (Davis, 2005) and Spanish frequencies < 15 words per million (Alameda & Cuetos, 1995) were selected. Because previous research has demonstrated that word length does have an effect on participants' ability to learn new words (Elgort, 2011), all target words contained between 3 and 7 letters in English, and between 4 and 9 letters in Spanish.

Cognates were selected so that the target word could be easily identified across languages for the mixed language context conditions. Short words were selected to minimize verbal working memory load. Cognate status was verified by the degree of orthographic similarity in both languages, and computed using the algorithm from Van Orden (1987). The algorithm computes an orthographic similarity score based on the ratio between graphemic similarities of two words relative to the graphemic similarity of one word to itself. Graphemic similarity is computed by a formula that takes into account: 1) the number of adjacent letter pairs two words share in order, 2) the number of adjacent letter pairs two words share in reversed order, 3) the number of letters shared, 4) the average word length of each word, 5) the ratio of the word length of the shorter word to the longer word, and 6) whether initial and final letters are the same in both words. All target words in the original pool had an orthographic overlap rating of greater than .30 (M = .70).

A norming experiment was conducted to verify that the selected pool of words were short and simple enough to be learned at an acceptable rate after 4 exposures. Thirty participants studied lists of 53 words in four repeated study blocks. Each word was presented in isolation, and the order of items was randomized within each block. Immediately following the last study block, participants completed a recognition memory test. During the testing phase, participants were presented 106 words in isolation on the computer screen, half of which were presented in the study session, and half of which were new items. They were instructed to decide as quickly and as accurately as possible whether or not that word was presented in the study session. The dependent variable was discrimination score (d'). Words that had d' of less than 1 (n = 16) were removed from the pool.

Critical sentence stimuli. A total stimulus set of 480 biasing sentences (four sentences per critical word) were constructed and normed to ensure that they adequately conveyed the meaning of the critical word. Thirty participants were presented 53 multiple choice items on a computer. Each item consisted of a target word presented in four sentence contexts intended to clearly convey its meaning. Participants were instructed to select the correct definition of the target word from a list of four definitions. To be included in the final stimulus set, critical sentences had to have accuracy rates greater than .75. Of the sentences meeting that criterion, the 120 critical words with the highest discrimination ratings in the stimulus word norming were chosen for inclusion in the present study. Sentences were translated into Spanish for the Spanish and mixed sentence conditions. The average length of the sentences was 15.75 (SD = 3.9) words for English sentences and 16.98 (SD = 4.5) words in Spanish sentences (See Table 2.2 for an example stimuli set for one item across all experimental conditions).

Sentences were counterbalanced across 24 experimental running lists based on a Latin square. For each list, 60 words were presented in the study phase (10 words per study condition), while the other 60 served as non-studied control items for the testing phase. Participants were randomly assigned to experimental lists.

Table 2.2: Example stimuli set for one item across all learning conditions

Repeated-English			
The players were making **<u>caustic</u>** comments, such as "this guy is awesome," about the horrible referee all game.	The players were making **<u>caustic</u>** comments, such as "this guy is awesome," about the horrible referee all game.	The players were making **<u>caustic</u>** comments, such as "this guy is awesome," about the horrible referee all game.	The players were making **<u>caustic</u>** comments, such as "this guy is awesome," about the horrible referee all game.
Varied-English			
The players were making **<u>caustic</u>** comments, such as "this guy is awesome," about the horrible referee all game	One of the most **<u>caustic</u>** comments from the evolutionary scientist was his stated belief in the creationism story in the Bible.	Her **<u>caustic</u>** comments towards her roommates were always sarcastic with a hint of truth in them.	The angry comic was always making **<u>caustic</u>** statements about society as a great place.
Repeated-Spanish			
Los jugadores estaban haciendo comentarios **<u>cáusticos</u>**, tales como "este hombre es increíble", sobre el árbitro horrible por todo el partido.	Los jugadores estaban haciendo comentarios **<u>cáusticos</u>**, tales como "este hombre es increíble", sobre el árbitro horrible por todo el partido.	Los jugadores estaban haciendo comentarios **<u>cáusticos</u>**, tales como "este hombre es increíble", sobre el árbitro horrible por todo el partido.	Los jugadores estaban haciendo comentarios **<u>cáusticos</u>**, tales como "este hombre es increíble", sobre el árbitro horrible por todo el partido.
Varied-Spanish			
Los jugadores estaban haciendo comentarios **<u>cáusticos</u>**, tales como "este hombre es increíble", sobre el árbitro	Uno de los comentarios más **<u>cáusticos</u>** desde el científico evolucionista fue su creencia declarada en la historia del	Sus comentarios **<u>cáusticos</u>** hacia sus compañeros eran siempre sarcásticos con una pizca de verdad en ellos.	El cómic enojado siempre estaba haciendo declaraciones **<u>cáusticas</u>** sobre la sociedad como un gran lugar.

horrible por todo el partido	creacionismo en la Biblia.		
Repeated-Mixed language			
The players were making **caustic** comments, such as "this guy is awesome," about the horrible referee all game.	Los jugadores estaban haciendo comentarios **cáusticos**, tales como "este hombre es increíble", sobre el árbitro horrible por todo el partido.	The players were making **caustic** comments, such as "this guy is awesome," about the horrible referee all game.	Los jugadores estaban haciendo comentarios **cáusticos**, tales como "este hombre es increíble", sobre el árbitro horrible por todo el partido.
Varied-Mixed Language			
The players were making **caustic** comments, such as "this guy is awesome," about the horrible referee all game.	Uno de los comentarios más **cáusticos** desde el científico evolucionista fue su creencia declarada en la historia del creacionismo en la Biblia.	Her **caustic** comments towards her roommates were always sarcastic with a hint of truth in them.	El cómic enojado siempre estaba haciendo declaraciones **cáusticas** sobre la sociedad como un gran lugar.

ESPADA. (Francis & Strobach, 2013). The ESPADA (English-Spanish Proficiency and Dominance Assessment) is a self-report questionnaire on participants' language usage. The questionnaire contains items that assess proficiency levels across four domains (reading, writing, speaking, and listening), as well as items that assess frequency of use, and age of acquisition information for both English and Spanish.

Woodcock-Muñoz Language Survey Revised (WMLS-R) (Woodcock et al., 2005). The WMLS-R is a standardized battery of tests used to assess proficiency in English and Spanish. In the present study, Test 1 (Picture Vocabulary) and Test 7 (Passage Comprehension) were administered to assess both oral and reading proficiency in English and Spanish. For the Picture Vocabulary subtest, participants are asked to name pictures aloud. For the Passage Comprehension subtest, participants read sentences and short passages with one word missing,

and are asked to say the missing word aloud. Standard scoring criteria were used for both subtests. Age-equivalency scores for each test were calculated via the computer software provided with the test.

2.1.3 Design

Experiment 1 had a 3 (language of context) by 2 (semantic variability of contexts) within-subjects design. Participants studied unfamiliar words either exclusively in L1, L2 or in both languages. Also, words were either studied in the same sematic context (identical sentences for the within language condition, translation of sentences for language mixed condition) or in varied semantic contexts. The dependent variable for the meaning generation task was accuracy proportions. The dependent variables for the semantic relatedness task were discrimination (d') scores, criterion scores (λ) and bias scores ($\log\beta$), and response times (RT).

2.1.4 Procedure

After completing the informed consent procedures, participants completed the ESPADA language history questionnaire (Francis & Strobach, 2013), and the picture vocabulary and reading comprehension subtests in both English and Spanish versions of the Woodcock-Muñoz Language Survey-Revised (Woodcock et al., 2005). In order to complete the experiment, participants had to have an age-equivalency score of at least 6 years old in both English and Spanish on both subtests.

Study Phase. Participants read a list of 60 words in sentences, with 10 words for each critical condition (L1 repeated contexts, L1 varied contexts, L2 repeated contexts, L2 varied contexts, mixed language repeated contexts, mixed language varied contexts). Each word was presented four times across four different blocks within the list. Trial order was randomized within each block. Words assigned to the repeated semantic context conditions (n = 30) were

presented in the same sentence context across all four blocks. One-third of these were always in the L1, another one-third was always in the L2, and the final third in both languages. Words assigned to the varied semantic context condition (n = 30) were presented in four different sentence contexts across the four blocks. One-third of these were always in the L1, another one-third always in the L2, and the final one-third in both languages. Twenty-four experimental lists were created such that all words were presented in all six conditions, with assignment to conditions counterbalanced across lists.

Target words were presented above each sentence, as well as embedded in the sentence, for each trial. Participants were instructed to read each sentence with the goal of extracting the meaning of the target word, as they would be tested on the meanings of those words following the study session. The target word and sentence appeared on the screen until participants pressed the spacebar to move on to the next trial. No constraints were placed on the amount of time participants could study each sentence.

Testing Phase. Immediately following the study session, participants completed two tasks during the testing phase, a meaning generation task and a semantic relatedness judgment task. The order of the testing tasks was counterbalanced across participants.

Meaning generation task. All 60 words that were presented during the study phase were presented at test. There were no filler or control items presented in this task. Words were presented one at a time on the screen following a fixation cross. Participants were instructed to say aloud a definition or synonym for each word that appeared on the screen while their responses were recorded. Participants pressed the spacebar to move to the next trial. Stimuli consisted of the same 60 target words presented in the study session. Responses were then transcribed and responses were rated separately by two experimenters as correct ('1') or

incorrect (‘0’). Raters were blind to experimental condition when coding for correct responses. Inter-rater reliability was assessed by the proportion of items that were agreed upon by both raters. Inter-rater reliability was .90 across all items. Discrepancies in ratings were discussed until both raters agreed on a correct rating.

Semantic relatedness task. Two words were presently on the screen simultaneously. Each trial began with a fixation cross appearing in the middle of the screen. Target words were paired with either a semantically related word or an unrelated word. Participants were instructed to press the ‘YES’ key if the words were related in meaning, or press the ‘NO’ key if the words were not related in meaning. The word pairs disappeared when a response was made. Participants were instructed to press the middle key on a button box to initiate the next trial. Each participant completed 120 trials. Trial order was randomized for each participant. For half of the trials, one of the words in the word pair was a word from the study phase. The other 60 trials consisted of words from the original pool of target stimuli, but not presented to the participant in the study phase. Half of the word pairs were semantically related in meaning (“YES” trials; *caustic-sarcastic*), and half were unrelated in meaning (“NO” trials; *caustic-invest*). The semantically-related meanings were created by the experimenter, and were high frequency synonyms or related words. As an additional control, half of the trials consisted of English word pairings, while the other half consisted of Spanish word pairings.

2.1.5 Apparatus

The computer portion of the experiment was presented on a Dell desktop computer using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). Responses were entered using an E-Prime button box.

2.1.6 Analysis

The meaning generation data were analyzed using logistic generalized linear mixed effects (LME) models within the lme4 package (Bates, 2007; Bates & Sarkar, 2006) of R (version 2.13.1; Baayen, 2008; Baayen et al., 2008; R Development Core Team, 2010). According to previous research, models should use a maximal random effects structure, including random slopes for items and participants, as well as by-item and by-participant random slope adjustments (Barr, Levy, Scheepers, & Tily, 2013). However, in practice, models with full random effects structures frequently do not converge (Whitford & Titone, 2017; Lauro & Schwartz, under review). Therefore, a model comparison approach was taken in which the simplest model always included the following fixed effects factors: language of study (English, Spanish, or Mixed), semantic context at study (varied sentences vs. repeated sentences), and their interaction. The simplest model also contained random intercepts for items and participants. By-participant and by-item random slope adjustments were incrementally added to the model to see whether the model fit was improved. Model fit was assessed by comparing log-likelihood values of the simpler and more complex model using a χ^2distribution. If the χ^2 value for the improved fit of the more complex model was significant, that model was used for analysis. This approach was taken so that each model represented the maximal random effects structure that is justified by comparing a more specified model to a more general model.

For the semantic relatedness task, an equal-variance signal detection analysis was used, and detection (d'), criterion (λ), and bias ($\log\beta$) scores were calculated for each participant (Wickens, 2002). The detection measure d' indicates the ability to discriminate signal trials (here, related trials) from noise trials (here, unrelated trials) and takes into account YES responses to related trials (hits) and YES responses to unrelated trials (false alarms). The criterion score λ indicates the threshold or amount of evidence needed to respond YES. It is

assumed that the same criterion is used for both signal and noise trials, because the participant does not know a priori which kind of trial is being presented. On a given trial, when the amount of evidence is larger than the criterion, participants respond YES. If the amount of evidence is less than the criterion, participants respond NO. Response bias (relative to the setting the criterion to maximize accuracy) is measured using $\log\beta$. For this measure, negative values indicates a bias to respond YES, while positive values indicate a bias to respond NO. These dependent variables were each submitted to a 3 (study language) x 2 (study sentence context) repeated-measures ANOVA. RT data for correct responses were analyzed using the same model comparison approach to LME modeling as the meaning generation data.

2.2 Results

2.2.1 Meaning Generation Task

To examine the effects of the independent variables on meaning generation accuracy, a logistic LME model was conducted. A model containing by-participant random slope adjustments for study sentence context significantly improved the model fit from the simplest model, containing only the fixed factors and random intercepts for items and participants, $\chi^2(3) = 8.11$, $p < 0.04$. Including by-item random slope adjustments for study language did not significantly improve the model fit of the simplest model, $\chi^2(6) = 3.496$, $p = 0.74$, nor did including by-item random slope adjustments for study sentence, $\chi^2(3) = 0.68$, $p = 0.88$. A model including by-participant random slope adjustments for study language failed to converge. Therefore, the model with the best fitting random effects structure for this analysis included all fixed factors, random intercepts for participants and items, and by-participant random slope adjustments for study sentence context.

For the purposes of the LME models, the effect of study language was coded using two contrast components, one that compared L1 and L2 conditions and one that compared single-language and mixed-language conditions. The first contrast on study language was significant (β = -0.28, *SE* = 0.05, z = -6.165, $p < .001$), indicating that participants had a higher proportion of accurate responses for words studied exclusively in L1 (*M* = .48) compared to words studied exclusively in L2 (*M* = .38). The second contrast on study language was not significant (single language *M* = .43, mixed language *M* = .43) that the proportions were not significantly different for words studied exclusively in one language (*M* = .43).

There was a significant effect of study sentence condition (β = 0.21, *SE* = 0.09 z = 2.40, p = .017). Participants had a higher proportion of accurate responses to words studied in varied sentences (*M* = .45) compared to words studied in repeated sentences (*M* = .41).

There was no interaction of the first contrast on study language with study sentence conditions (β = -0.043, *SE* = 0.09, z = -0.474), p = .64), indicating that the magnitude of the advantage for words studied in varied sentences was equivalent whether they were studied exclusively in L1 or exclusively in L2. However, there was a significant interaction between the second contrast on study language and study sentence conditions (β = .015, *SE* = 0.05, z = -2.97, $p < .003$), indicating that the effects of sentence context variation were observed for words studied exclusively in one language but not for words that were studied in both languages. Effect sizes, standard errors, z-scores and significance levels for the logistic LME model are reported in Table 2.3. Mean accuracy proportions are reported in Table 2.4.

Table 2.3: Effect size, standard error and z-values for the logistic LME conducted on accuracy for the meaning generation task.

	Accuracy		
Fixed effects	**B**	**SE**	**z**
Study Language[1]			
L1 vs L2	-0.284	.05	-6.17***
Single language vs both languages	0.003	.03	0.10
Study Sentence[2]	0.211	.09	2.40*
Interaction			
Study Language (L1 vs L2) x Study Sentence (repeated vs varied)	-0.043	.09	-0.474
Study Language (single vs both) x Study Sentence (repeated vs varied)	0.152	.05	2.97**
Intercept	-0.406	.16	-2.56*

	Variance	
Random effects	Intercept	Slope (Sentence)
Participants	0.583	0.164
Items	1.193	--
Residual (Median)	-0.295	

[1]contrasts were deviation coded (L1 vs. L2, single language vs. both languages). [2]contrasts were deviation coded (Repeated vs. Varied). *** $p < .001$ ** $p < .01$ * $p < .05$.

Table 2.4: Means (SE) for proportion of correct responses in the meaning generation task

	Accuracy Proportions	
	Repeated	Varied
L1	.45	.52
	(.02)	(.02)
L2	.35	.41
	(.02)	(.02)
Mixed	.44	.43

(.02)	(.02)

2.2.2 Semantic Relatedness Task

To make the semantic-relatedness task meaningful, half of the trials consisted of non-studied word-pairs. Discrimination scores were obtained for each participant in each cell of the design. Preliminary *t* tests revealed that discrimination scores for word pairs in which the target word was studied were larger than discrimination scores for word pairs that did not include any studied words, $t(71) = 10.38$, $p < .001$. Specifically, discrimination scores for word pairs that did not include any studied items were 0.70 when tested in L1 and 0.38 when tested in L2. Although, these values were significantly greater than chance, indicating at least some prior knowledge of non-studied words, $t(71) = 9.06$, $p < .001$, and $t(71) = 5.21$, $p < .001$ respectively. Additionally, RTs were faster to word-pairs that included a studied word (M = 2826 ms) than word-pairs that did not include any studied word (M = 3143 ms), $t(71) = -5.95$, $p < .001$. Specifically, the average RT for non-studied word-pairs was 3077 ms in L1, and 3208 ms in L2.

Discrimination Scores. A 3 (study language) x 2 (sentence context) repeated-measures ANOVA was conducted on d' discrimination scores. The analysis on d' revealed a significant effect of study language, $F(2, 140) = 7.23$, MSE = 4.73, $p < .02$. Follow-up pairwise comparisons revealed that discrimination scores for critical words studied and tested exclusively in L1 (M = 1.67) were greater than critical words studied and tested exclusively in L2 (M = 1.32), $p < .001$; and critical words studied and tested in both languages (M = 1.41), $p < .01$. There was no significant difference in discrimination scores between critical words studied and tested in L2 or in both languages. There was no significant main effect of study sentence context (repeated M = 1.45; varied M = 1.48). The interaction between language and sentence study

context was not significant (means and standard errors for discrimination scores are reported in Table 2.5).

Criterion Scores. A 3 (study language) x 2 (sentence context) repeated-measures ANOVA was conducted on criterion scores. The analysis on criterion scores (λ) yielded a significant effect of study language $F(2, 70) = 3.46$, MSE = 1.66, $p < .04$. Pairwise comparisons revealed that criterion scores were significantly higher for word pairs in which the critical word was studied exclusively in L1 ($M = .63$) compared to word pairs in which the critical word was studied exclusively in L2 ($M = .42$), $p < .005$. There were no significant differences between criterion scores for word-pairs in which the critical word was studied in both languages and when the critical word was studied exclusively in L1 or exclusively in L2. There was no significant effect of study sentence context, nor was there a significant interaction. Means and standard errors for lambda scores are reported in Table 2.5.

Bias Scores. A 3 (study language) x 2 (sentence context) repeated-measures ANOVA was conducted on logβ bias scores. There was no significant main effect of study language or study sentence context. However, the interaction was significant, $F(2, 70) = 5.26$, MSE = 6.82, $p < .006$. Specifically, there was a greater YES bias for words studied in varied sentence contexts in the L1 pure condition (repeated sentences logβ = -.71; varied sentences logβ = -.88) and L2 pure condition (repeated sentences logβ = -.34; varied sentences logβ = -.69). However, in the mixed-language condition, there was less of a YES bias for words studied in varied sentences (logβ = -.25) than repeated sentences (logβ = -.75). Means and standard errors for logβ are reported in Table 2.5.

Table 2.5: Mean (SE) Discrimination, Criterion, and Bias scores for the Semantic Relatedness Task

	d'		**Criterion Scores (λ)**		**Bias Scores (Logβ)**	
	Repeated	Varied	Repeated	Varied	Repeated	Varied
L1	1.67	1.67	.65	.62	-.71	-.88
	(.13)	(.12)	(.09)	(.08)	(.20)	(.20)
L2	1.26	1.37	.42	.41	-.34	-.69
	(.12)	(.12)	(.09)	(.08)	(.12)	(.16)
Mixed	1.42	1.41	.55	.53	-.74	-.25
	(.13)	(.12)	(.09)	(.10)	(.20)	(.12)

Response Times. Response-time analyses were conducted on trials with correct responses only. Because there were no time constraints placed on participant responses, outliers (+/- 2 standard deviations from the mean) were removed. This resulted in the exclusion of 4% of the RT data.

To analyze the effects of the independent variables on response times (RTs), an LME model analysis was conducted. Fixed effects factors were study language (L1, L2, or Mixed), study sentence (repeated vs. varied), and their interaction. The simplest model contained all fixed effects factors and random intercepts for participants and items. Models including by-items and/or by-participant random slope adjustments for study sentence did not converge. A model including by-item random slope adjustments for study language did not significantly improve the model fit, $\chi^2(6) = 6.63$, $p = .36$. A model including by-participant random slope adjustments for study language did failed to converged. Therefore, the final model used for

analysis of RTs was the simplest model, containing fixed factors and random intercepts for participants and items.

For the purposes of the LME models, the effect of study language was coded using two contrast components, one that compared L1 and L2 conditions and one that compared single-language and mixed-language conditions. The first contrast on study language was significant (β = 142.14, *SE* = 32.02, *t* = 4.40, *p* < .001), indicating that word pairs in which the critical words were studied exclusively in L1 (*M* = 2618 ms) were faster than responses to word pairs in which the critical word was studied exclusively in L2 (*M* = 2899 ms). The second contrast on study language was not significant β = -18.67, *SE* = 18.35, *t* = -1.02, *p* = .31), indicating no difference in RTs for word pairs in which the critical word was studied in one language (either exclusively in L1 or exclusively in L2) (*M* = 2759 ms) and word pairs in which the critical word was studied in both languages (*M* = 2791 ms).

There was no significant effect of study sentence (varied vs repeated) context was not significant. However, there was a significant interaction between the first study language contrast and study sentence context (β = 147.26, *SE* = 63.86, *t* = 2.31, *p* < .02). RTs for correct responses to word pairs in which the critical word was studied in L1 were faster when they were studied in four varied sentences (*M* = 2704 ms) compared to word pairs in which the critical word was studied in L1, but in the same repeated sentence (*M* = 2704 ms). RTs for correct responses to word-pairs in which the critical word was studied in L2 were faster when the study sentences were repeated (*M* = 2825 ms) compared to varied sentences (*M* = 2971 ms). RTs for correct responses to word-pairs in which the critical word was studied in both languages were similar when the critical word was studied in repeated sentences (*M* = 2799 ms) or varied

sentences (*M* = 2783 ms) (see Table 2.6 for effect sizes, standard errors, and significance of the LME model for RTs of correct trials on the semantic relatedness task).

Table 2.6: Effect size, standard error, and t-values for LME model for the semantic relatedness task

	Response Time		
Fixed effects	**B**	**SE**	**t**
Study Language[1]			
L1 vs L2	142.139	32.016	4.44***
Single language vs both languages	-18.671	18.349	-1.02
Study Sentence[2]	-16.049	51.728	-0.31
Interaction			
Study Language (L1 vs L2) x Study Sentence (repeated vs varied)	147.260	63.863	2.31*
Study Language (single vs both) x Study Sentence (repeated vs varied)	-5.692	36.709	-0.16
Intercept	2783.089	73.45	37.89***
	Variance		
Random effects	Intercept	Slope (Sentence)	
Participants	279467	--	
Items	97283	--	
Residual (Median)	2000858		

[1]contrasts were deviation coded (L1 vs. L2 vs. Mixed). [2]contrasts were deviation coded (Repeated vs. Varied). *** $p < .001$; * $p < .05$.

2.3 Discussion

Consistent with previous research, an advantage in recall was observed for words that were presented in varied semantic contexts relative to repeated contexts. This was evident in the proportion of accurate meanings generated for words encountered in L1 pure and L2 pure study conditions. To our knowledge, this is the first study to explicitly test the advantage of varied

semantic contexts in bilinguals L2. The magnitude of the benefit for varied semantic contexts was similar in L1 and L2, suggesting that the distinctiveness of the memory trace was similarly enhanced, regardless of whether comprehending sentences in a stronger or weaker language.

We had also predicted that there would be an advantage for studying novel words in mixed languages versus in language pure environments. This prediction was not supported, as overall performance for words studied in both languages was not superior to words studied purely in L1 or purely in L2.

Accuracy was lowest for words encountered in the L2 and in the same semantic context. This is likely due to the compounding of two disadvantages, that is, 1) encountering words in a weaker language, and 2) no semantic contextual variation. This suggests that, in the absence of distinctive semantic cues, distinctive language cues enhance the memory trace.

One unexpected finding was that there was no benefit of semantic variation in the mixed language study condition. This finding can be understood by considering how language mixed conditions reduced the salience of semantic cues while at the same time enhanced the salience of language context. Specifically, in language mixed conditions, semantic information from the L2 sentences may have not been as deeply encoded- thus reducing the overall semantic salience of the varied contexts. At the same time, having repeated sentences in two different languages enhanced the memory trace by creating language-context cues. That is, in the absence of semantic variation, encountering a word in two languages yielded more accurate responses than only encountering a word in L2. Perhaps subsequent word encounters more strongly activate previous L1 traces relative to L2.

The results of the semantic relatedness task did not provide direct evidence for an advantage in semantic contextual variation at study. Two factors may have contributed to the

lack of support for semantic context variation. The first is the nature of the task. In the semantic relatedness task, all related items were high frequency synonyms of the studied word. The synonym at test may have acted as a more powerful retrieval cue, overriding the benefit of more distinctive traces generated at study. A second, related issue is that fact that no time constraints were placed on participants' responses. Even eliminating outliers, the range in response times was quite large (10 seconds). To maximize the potential that context-independent knowledge of a word results in an advantage in recognition memory, and to minimize variability in response times, Experiment 2 placed a five-second time constraint on responses.

However, there was some evidence that the language in which a word was presented did interact with semantic context. RTs were faster for words studied in varied semantic contexts, but only when words were studied exclusively in L1. This finding is consistent with the advantage of semantic contextual variation observed in monolingual studies (Balass, 2011; Bolger et al., 2008) and provides supporting evidence that L1 traces of encounters with words may have been stronger than L2 traces. When operating in their dominant language, participants accessed and retrieved semantic knowledge of the critical word more quickly because more semantic associations were formed while studying words in varying semantic contexts. When operating in a weaker language, there was no advantage in RTs for semantic variation at study.

Additionally, an interesting interaction was observed in the analysis of bias scores of the semantic relatedness task. There was a significantly smaller YES bias when words were studied in both languages, and in varying semantic contexts. Although there was no difference observed in the accuracy of responses, perhaps the combined effect of language distinctiveness and semantic distinctiveness provided more contextual cues when encountering the words again during the testing phase, limiting biased responses.

Chapter 3: Experiment 2

Experiment 1 replicated previous research on the advantage of semantic contextual variation in visual presentation of novel words. Though, the predicted effect of language distinctiveness was not strongly supported. It might be that any potential benefit of language distinctiveness across encounters with novel words is attenuated in written form for languages that share the same script. However, even languages that share the same script have unique phonetic realizations for words. For example, the English-Spanish cognate, *base*, has identical orthography, but distinct pronunciations. In Experiment 2, we attempted to further the distinctiveness across language contexts by presenting novel words auditorily. The logic behind this is that English and Spanish, although distinct, share the same orthographic script. However, phonologically, the two languages are more distinct. Thus, it might be the case that mixing languages at study might enhance the previously observed effect of semantic variation.

3.1 Method

3.1.1 Participants

Seventy-two highly proficient Spanish-English bilingual speaking undergraduates (48 females, 24 males) from the University of Texas at El Paso participated for course credit. The median age of participants was 20. The majority of the participants learned Spanish first. The mean AoA for English was 7 (range 1-15), while the mean AoA for Spanish was 2 (range 0-10).

Language Proficiency. Participants' language proficiency was assessed using the same subjective (ESPADA: Francis & Strobach, 2013) and objective (WMLS-R: Woodcock-Muñoz: Woodcock et al., 2005) measures in Experiment 1 (see materials section for Experiment 1 for a description of these instruments). Participants were classified as being English-dominant or Spanish-dominant based on their relative Picture Vocabulary scores. Based on this criterion, 39

participants were classified as English dominant and 33 were Spanish dominant (see Table 2.1 for summary information on participants).

The English-dominant group had slightly higher Picture Vocabulary scores in their L1 relative to the L1 scores of the Spanish-dominant group (M = 15.4 and M = 13.0, respectively), $t(70) = 2.88$, $p = .005$. There was no significant difference between the language dominance groups in L1 passage comprehension, $t(70) = 1.09$, $p = .28$. The small difference in general L1 proficiency, as measured by the Picture Vocabulary subtest, may be due to the fact that all participants were receiving formal education in English, so even the Spanish-dominant group was exposed to more English.

Spanish-dominant participants' L2 reading comprehension age equivalencies in English (M = 13.32) was higher than English-dominant participants' L2 reading comprehension age equivalencies in Spanish (M = 10.36). However, there was no significant difference between language dominance groups in L2 passage comprehension, $t(70) = 1.61$, $p = 0.11$.

Self-ratings on the subjective measure aligned with the classification of English versus Spanish dominant. English dominant participants rated themselves as having better proficiency in English than Spanish, whereas Spanish dominant participants rated themselves as having better proficiency in Spanish than in English.

3.1.2 Materials

The materials were the same as in Experiment 1 except that all words and sentences were presented auditorily rather than visually. Sentences and critical words were recorded by a female, native speaker of English and Spanish using the audio recording software, PRAAT (Boersma & Weenink, 2017). Audio files were then converted to .wav files and uploaded to the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

3.1.3 Design

Experiment 2 had a 3 (language of context) by 2 (semantic variability of contexts) within-subjects design. Participants studied unfamiliar words auditorily either exclusively in L1, L2 or in both languages. Also, words were either studied in the same sematic context (identical sentences for the within language condition, translation of sentences for language mixed condition) or in varied semantic contexts. The dependent variable for the meaning generation task was accuracy proportions. The dependent variables for the semantic relatedness task were discrimination (d') scores, criterion scores (λ) and bias scores (logβ), and response times (RT).

3.1.4 Procedure

Participants completed the same questionnaire and language assessments as in Experiment 1.

Study Phase. Participants listened to a list of 60 words in sentences, with 10 words for each critical condition (L1 repeated contexts, L1 varied contexts, L2 repeated contexts, L2 varied contexts, mixed language repeated contexts, mixed language varied contexts). Each word was presented four times across four different blocks within the list. Within each block, trial order was randomized. Words assigned to the repeated semantic context conditions (n = 30) were presented in the same sentence context across all four blocks. One-third of these were always in the L1, another one-third was always in the L2, and the final third in both languages. Words assigned to the varied semantic context condition (n = 30) were presented in four different sentence contexts across the four blocks. One-third of these were always in the L1, another one-third always in the L2, and the final one-third in both languages. Twenty-four experimental lists were created such that all words were presented in all six conditions, with assignment to conditions counterbalanced across lists.

Target words were heard before each sentence. Participants were instructed to listen to each sentence with the goal of extracting the meaning of the target word, as they would be tested on those words following the study session. Although this was a self-paced task (participants pressed the spacebar to move on to subsequent trials), participants could only listen to each sentence once before moving on to the next trial.

Testing Phase. Immediately following the study session, participants completed two separate tasks at test, a meaning generation task and a semantic relatedness judgment task. The order of the tasks was counterbalanced across participants.

Meaning generation task. All 60 words that were presented during the study phase were presented at test. There were no filler or control items presented in this task. Words were presented one at a time, auditorily, following a fixation cross on the computer screen. Participants were instructed to say aloud a definition or synonym for each word they heard while their responses were recorded. To move to the next trial, participants pressed the middle key on a button box. Responses were then transcribed and responses were rated separately by two experimenters as correct ('1') or incorrect ('0'). Inter-rater reliability was assessed by the percentage of items that were agreed upon by both raters. Inter-rater reliability was 0.93 across all items. Discrepancies in ratings were discussed until both raters agreed on a correct rating.

Semantic relatedness task. Participants listened to 120 word pairs. Half of the word pairs contained a critical word that was studied in the study phase. The other 60 trials consisted of words from the original pool of target stimuli, but not presented to the participant in the study phase. For word pairs in which a studied word was included, the studied word was always heard first, followed by a semantically related or unrelated word. Following the auditory presentation of the second word (for studied and non-studied word pair trials), participants were instructed to

press the ‘YES’ key if the words were related in meaning, or press the ‘NO’ key if the words were not related in meaning within five seconds. Each participant completed 120 trials. Half of the word pairs were semantically related in meaning (“YES” trials: *caustic-sarcastic*), and half were unrelated in meaning (“NO” trials; *caustic-invest*). The semantically-related meanings were created by the experimenter, and were high frequency synonyms or related words. As an additional control, half of the trials consisted of English word pairings, while the other half consisted of Spanish word pairings. Additionally, we attempted to reduce variability of correct responses observed in Experiment 1 by implementing a 5 second time limit on responses.

3.1.5 Apparatus

The computer portion of the experiment was presented on a Dell desktop computer using Eprime software (Psychology Software Tools, Pittsburgh, PA). Responses were entered using an Eprime button box. Audio recordings were presented with RadioShack headphones that covered the entire ear. Participants could adjust the volume to their liking.

3.1.6 Analyses

The data for the meaning generation task were analyzed using the same model comparison approach to LME models as Experiment 1. The data were analyzed within the lme4 package (Bates, 2007; Bates & Sarkar, 2006) of R (version 2.13.1; Baayen, 2008; Baayen et al., 2008; R Development Core Team, 2010).

3.2 Results

3.2.1 Meaning Generation Task

To examine the effects of the independent variables on meaning generation accuracy, a logistic LME model was conducted. Including by-item random slope adjustments for study language did not significantly improve the model fit compared to the simplest model, χ^2 (6) =

1.091, $p = 0.98$), nor did including by-participant random slope adjustments for study language, χ^2 (6) = 11.16, $p = .08$). A model including by-participant random slope adjustments for study sentence did not significantly improve the simplest model, χ^2 (3) = 0.24, $p = .97$). The model including by--item random slope adjustments for study sentence context failed to converge. Therefore, the simplest model, including all fixed factors (study language and study sentence context) and random intercepts for participants and items, was adopted for the final analysis.

For the purposes of the LME models, the effect of study language was coded using two contrast components, one that compared L1 and L2 conditions and one that compared single-language and mixed-language conditions. The first contrast on study language was significant ($\beta = -0.36$, $SE = 0.05$ z = -7.278, $p < .001$), indicating a higher proportion of accurate responses for words studied exclusively in L1 ($M = .33$) compared to words studied exclusively in L2 ($M = .23$). The second contrast on study language was not significant ($\beta = 0.016$, $SE = 0.03$ z = 0.561, $p = .57$). The proportion of accurate responses was not different for words studied exclusively in one language ($M = .28$) and words studied in both languages ($M = .27$).

There was also a significant effect of study sentence condition ($\beta = 0.20$, $SE = 0.08$ z = 2.47, $p < .013$). Participants had a higher proportion of accurate responses to words studied in varied sentences (.29) compared to words studied in repeated sentences (.26). There was no significant interaction between either study language contrast and study sentence context (effect sizes, standard errors, z-scores and significance levels for the logistic LME model are reported in Table 3.1 and means and standard deviations are reported in Table 3.2).

Table 3.1: Effect size, standard error and z-values for the logistic LME conducted for the Meaning Generation Task

	Accuracy		
Fixed effects	**B**	**SE**	**z**
Study Language[1]			
L1 vs L2	-0.360	.05	-7.278***
Single language vs both languages	0.016	.03	0.561
Study Sentence[2]	0.197	.08	2.479*
Interaction			
Study Language (L1 vs L2) x Study Sentence (repeated vs varied)	-0.034	.10	-0.342
Study Language (single vs both) x Study Sentence (repeated vs varied)	-0.050	.06	0.369
Intercept	-1.353	.16	-8.43***

Random effects	**Variance** Intercept	Slope
Participants	1.033	--
Items	1.113	--
Residual (Median)	-0.314	

[1]contrasts were deviation coded (L1 vs. L2: single language vs. mixed). [2]contrasts were deviation coded (Repeated vs. Varied). *** $p < .001$; * $p < .05$.

Table 3.2: Mean (SE) proportion of correct responses in the meaning generation task

	Accuracy Proportions	
	Repeated	Varied
L1	.31	.34
	(.02)	(.03)
L2	.22	.23
	(.02)	(.03)
Mixed	.24	.29
	(.02)	(.02)

3.2.2 Semantic Relatedness Task

Half of the trials in the semantic relatedness task consisted of non-studied word-pairs. Preliminary *t* tests revealed that discrimination scores for word pairs in which the critical word was studied (d' = 0.87) were larger than discrimination scores for word pairs that did not include any studied words (d' = 0.20), $t(71) = 8.66$, $p < .001$. The average discrimination score for non-studied word pairs heard in L1 was 0.21, while the average discrimination score for non-studied word pairs heard in L2 was 0.18. Although, these values were significantly greater than chance, indicating at least some prior knowledge of non-studied words, $t(71) = 3.43$, $p < .001$, and $t(71) = .025$, respectively. Additionally, RTs were faster to word pairs that included a studied word (M = 2215 ms) than word pairs that did not include any studied words (M = 2287), $t = -3.07$, p < .002. RTs to non-studied word pairs were faster in L1 (M = 2257 ms) than L2 (M = 2318 ms), cite statistic?.

Discrimination Scores. A 3 (study language) x 2 (sentence context) repeated-measures ANOVA was conducted on d' discrimination scores. The analysis on d' revealed a significant effect of study language, $F(2, 142) = 11.32$, MSE = 9.02, $p < .001$. Follow-up pairwise comparisons revealed discrimination scores for critical words studied and tested exclusively in L1 (M = 1.15) were greater than critical words studied and tested exclusively in L2 (M = 0.67), $p < .001$; and critical words studied and tested in both languages (d' = 0.79), $p < .002$. There was no significant difference in discrimination scores between critical words studied and tested in L2 and those studied in both languages. There was no significant main effect of study sentence context (repeated M = 0.90; varied M = 0.83). The interaction between language and sentence study context was not significant (means and standard errors for discrimination scores are reported in Table 3.3).

Criterion Scores. A 3 (study language) x 2 (sentence context) repeated-measures ANOVA was conducted on criterion scores. The analysis on criterion scores (λ) did not yield any significant main effects or interactions. Means and standard errors for criterion scores are reported in Table 3.3.

Bias Scores. The analysis on bias scores (logβ) yielded a significant main effect of study sentence context, $F(1, 71) = 5.30$, MSE = 1.691, $p = .024$. Critical words studied in repeated sentences elicited a greater YES bias (logβ = -0.142) than critical words studied in varied sentences (logβ = -0.017). There was no significant main effect of study language, nor was the interaction significant. Means and standard errors for logβ scores are reported in Table 3.3.

Table 3.3: Means (SE) for participant responses in each condition for the semantic relatedness task

	d'		**Criterion Scores (λ)**		**Bias Scores (Logβ)**	
	Repeated	Varied	Repeated	Varied	Repeated	Varied
L1	1.26	1.03	.44	.46	-.24	.004
	(.12)	(.13)	(.08)	(.10)	(.08)	(.08)
L2	0.64	0.70	.31	.34	-.03	-.07
	(.13)	(.11)	(.09)	(.08)	(.08)	(.07)
Mixed	0.80	0.77	.31	.41	-.16	.02
	(.13)	(.12)	(.08)	(.09)	(.06)	(.08)

Response Times. To analyze the effects of the independent variables on RTs, a model comparison approach to LME modeling was adopted. Fixed-effects factors were study language (L1, L2, or Mixed), study sentence context (repeated vs. varied), and their interaction. The simplest model also included random intercepts for participants and items. A model including by-item random slope adjustments for study language failed to converge; and including by-

participant random slope adjustments for study language did not significantly improve the model fit of the simplest model, χ^2 (6) = 0.24, p = .99. A model containing by-participant random slope adjustments for study sentence context failed to converge. A model including by-item random slope adjustments for study sentence context did significantly improve the model fit of the simplest model, χ^2 (3) = 12.52, p < .006. Models including by-participant and/or by-item random slope adjustments for study language in addition to by-item random slope adjustments for study sentence condition failed to converge. Therefore, the model used for final analysis of RTs contained all of the fixed factors, random intercepts for participants and items, and by-item random slope adjustments for study sentence context.

For the purposes of the LME models, the effect of study language was coded using two contrast components, one that compared L1 and L2 conditions and one that compared single-language and mixed-language conditions. The first contrast on study language was significant (β = 42.066, *SE* = 16.738, t = 2.51, p = .012), indicating that RTs for word pairs in which the critical words were studied exclusively in L1 (*M* = 2152 ms) were faster than responses to word pairs in which the critical word was studied exclusively in L2 (*M* = 2236 ms). The second contrast on study language was marginally significant (β = -17.639, *SE* = 9.771, t = -1.81, p = .07), indicating that RTs for word pairs in which the critical word was studied in one language (either exclusively in L1 or L2) (*M* = 2194 ms) were faster relative to when the critical word was studied in both languages (*M* = 2264 ms) (β = -17.639, *SE* = 9.771, t = -1.05, p = .07).

There was no significant effect of study sentence context. There were no significant interactions among the fixed factors (See Table 3.4 for effect sizes, standard errors, and t-values of the LME model conducted on RTs).

Table 3.4: Effect size, standard error and t-values for LME models for the semantic relatedness task

	Response Time		
Fixed effects	**B**	**SE**	**t**
Study Language[1]			
L1 vs L2	42.066	16.738	2.513*
Single language vs both languages	-17.639	9.771	-1.805
Study Sentence[2]	13.034	33.515	0.389
Interaction			
Study Language (L1 vs L2) x Study Sentence (repeated vs varied)	36.364	33.511	1.085
Study Language (single vs both) x Study Sentence (repeated vs varied)	-28.566	19.562	-1.460
Intercept	2218.849	38.193	58.10***

	Variance	
Random effects	Intercept	Slope (Sentence)
Participants	84054	--
Items	9893	44217
Residual (Median)	513733	

[1]contrasts were deviation coded (L1 vs. L2; single language vs. mixed). [2]contrasts were deviation coded (Repeated vs. Varied). *** $p < .001$; * $p < .05$.

3.2.3 Cross-Experimental Comparisons

We had originally predicted that language-specific phonological cues would lead to a greater degree of linguistic distinctiveness between languages, and therefore language contextual effects would be stronger for words presented auditorily compared to words presented visually. Evidence comparing the impact of modality on memory for words presented visually and auditorily have been mixed. Some studies have found an advantage for visually presented stimuli (e.g. Nelson et al., 2005). One explanation of this pattern is activation of phonological information during reading (Perfetti et al., 1988). However, other studies have demonstrated that

orthographic codes can be activated for auditory input (e.g. Castles et al., 2003). The only study directly comparing word learning across modalities found an advantage for novel words (in sentences) presented visually (Nelson et al., 2005). However, the present study is the first to compare word learning across modalities in bilingual language contexts. Data from Experiments 1 and 2 were aggregated and re-analyzed adding the presentation modality (auditory or visual) context of the learning phase as a between-subjects factor.

Meaning Generation Task. The same model comparison approach to LME models, with the simplest model containing study language, study sentence context, and study modality as fixed factors and random intercepts for items and participants, was conducted on aggregated data across experiments 1 and 2. Models containing by-item and by-participant random slope adjustments for study language did not converge. A model including by-item random slope adjustments for study sentence did not significantly improve the model fit of the simplest model, χ^2 (3) = 1.68, p = .64. A model containing by-participant random slope adjustments for study sentence did not converge. Including by-item random slope adjustments for modality did significantly improve the model fit, χ^2 (3) = 24.64, p < .001.. Therefore, the model containing all fixed factors, random intercepts for participants and items, and by-item random slope adjustments for modality was analyzed.

For the purposes of the LME models, the effect of study language was coded using two contrast components, one that compared L1 and L2 conditions and one that compared single-language and mixed-language conditions. The first contrast on study language was significant (β = -0.32, *SE* = 0.03, z = -9.380, p < .001), indicating a greater proportion of accurate responses was reported for words studied in L1 (M = .41) than for words studied in L2 (M =.30). The second contrast on study language was not significant (β = 0.01, *SE* = 0.02, z = 0.513, p = .61).

The proportion of accurate responses was not significantly different for words studied exclusively in one language (*M* = .36) compared to words studied in both languages (*M* = .35).

There was also a significant effect of study sentence context (β = 0.22, *SE* = 0.05, *z* = 3.99, $p < .001$. Words studied in varied sentence contexts elicited a higher proportion of accurate responses (*M* = .37) than words studied in repeated sentence contexts (*M* = .33).

A significant difference between auditory and visual presentation modalities was also observed (β = 0.95, *SE* = 0.18, z = 5.144, p < .001). A greater proportion of accurate responses was reported for words studied visually (*M* = .43) than words studied auditorily (*M* =.27).

There were no significant two-way interactions. However, a three-way interaction between the second study language contrast (single language vs both languages), study sentence (repeated vs varied) and modality was observed (β = 0.19, *SE* = 0.08, z = 2.593, p < .009). The nature of this interaction is such that words studied in varied sentence contexts elicited a greater proportion of correct responses compared to words studied in the same repeated context when they were studied auditorily, and exclusively in L1, L2, or both languages. However, when words were studied visually, there was no difference in the proportion of correct responses to words studied in repeated or varied sentences when words were studied in both languages. When words were studied in only one language (whether in the L1 or the L2), there was a greater proportion of correct responses to words studied in varied sentences compared to the same repeated sentence. (See Table 3.5 for effect sizes, standard errors, and z-values for the logistic LME model conducted on the cross-experimental data, and Table 3.6 for means and SEs).

Table 3.5: Effect size, standard error and z-values for the logistic LME conducted on accuracy proportions for the meaning generation task aggregated across both experiments

	Accuracy		
Fixed effects	**B**	**SE**	**z**
Study Language[1]			
L1 vs L2	-0.316	0.034	-9.380***
Single language vs both languages	0.010	0.019	0.513
Study Sentence[2]	0.215	0.054	3.990***
Modality[3]	0.947	0.184	5.144***
Interaction	-0.036	0.067	-0.541
Study Language (L1 vs L2) x Study Sentence	0.049	0.038	1.281
Study Language (single vs both) x Study Sentence	0.089	0.067	1.316
Study Language (L1 vs L2) x Modality	-0.015	0.038	-0.407
Study Language (single vs both) x Modality	0.039	0.108	0.357
Study Sentence x Modality			
Study Language (L1 vs L2) x Study Sentence x Modality	0.019	0.133	0.143
Study Language (single vs both) x Study Sentence x Modality	0.197	0.076	2.593**
Intercept	-0.875	0.129	-6.768***

Random effects	**Variance** Intercept	Slope (Modality)
Participants	1.0098	--
Items	0.5443	0.2882
Residual (Median)	-0.3058	

[1]contrasts were deviation coded (L1 vs. L2, single language vs. mixed). [2]contrasts were deviation coded (Repeated vs. Varied). [3]contrasts were deviation coded (auditory vs visual).*** $p < .001$; * $p < .05$.

Table 3.6: Mean (SE) correct response proportions for auditory and visually presented stimuli across language conditions in the meaning generation task.

	Accuracy Proportions			
	Auditory		Visual	
	Repeated	Varied	Repeated	Varied
L1	.31 (.02)	.34 (.02)	.45 (.02)	.51 (.02)
L2	.22 (.02)	.24 (.02)	.35 (.02)	.41 (.02)
Mixed	.24 (.02)	.29 (.02)	.44 (.02)	.43 (.02)

Semantic Relatedness Task. To analyze the data across both experiments, a 3 (language) x 2 (sentence context) x 2 (modality) mixed-ANOVA was conducted, treating language and sentence contexts as within-subjects factors, and modality as a between-subjects factor. The dependent variables were discrimination scores, criterion scores, and bias scores.

Discrimination scores. The analysis on discrimination scores revealed the same significant effect of study language as observed in Experiments 1 and 2, $F(2, 282) = 18.45$, MSE = 0.73, $p < .001$. Follow-up pairwise comparisons revealed discrimination scores for critical words studied in the L1 pure condition ($M = 1.41$) were greater than critical words studied in L2 ($M = 0.99$), $p < .001$, and critical words studied in the mixed language condition ($M = 1.10$), $p < .001$. There was also a significant main effect of modality, $F(1, 141) = 29.08$, MSE = 2.64, $p < .001$. Critical words that were studied visually elicited greater discrimination scores ($M = 1.47$) than critical words studied auditorily ($M = 0.87$). There was no significant difference between critical words studied in L2 or mixed-language conditions. There was no significant effect of study sentence context. There were no significant interactions. Means and standard errors are reported in Table 3.7.

Criterion scores. The analysis on criterion scores (λ) revealed no significant main effects of study language or study sentence contexts. However, there was a significant effect of modality, $F(1, 141) = 4.09$, $p < .05$. Criterion scores were higher when the critical word in the word pair was studied visually ($M = .53$) compared to auditorily ($M = .38$). There were no significant interactions. Means and standard errors are reported in Table 3.7.

Bias Scores. The analysis on bias scores revealed no significant main effect of study language or study sentence. However, there was a significant effect of modality, $F(1, 141) = 28.83$, MSE = 2.02, $p < .001$, in which a greater YES bias was observed for words studied visually (logbeta = -.60) than for words studied auditorily (logβ = -.08). A significant interaction between study language and study context was observed, $F(2, 282) = 5.98$, MSE = .83, $p < .008$. When critical words were studied exclusively in L1, bias scores did not differ between the repeated (logβ = -.48) and varied (logβ = -.44) study sentence context conditions. When words were studied exclusively in L2, a greater YES bias was observed for critical words studied in varied sentences (logβ = -.38) than repeated sentences (logβ = -.19). However, words studied in both languages elicited a greater YES bias for words studied in repeated (logβ = -.45) compared to varied (logβ = -.16) sentence contexts.

Additionally, there was a significant 3-way interaction between study language, study sentence context, and presentation modality, $F(2, 282) = 3.31$, MSE = .83 $p < .04$. When words were studied in the auditory modality, a YES response bias was observed for words studied in repeated sentences, and minimal response bias was observed in words studied in varied sentences. This bias was greatest for words studied in the L1 pure condition (repeated logβ = -.24; varied logβ = -.004). In the mixed language condition, the same pattern was observed, but attenuated (repeated logβ = -0.16; varied logβ = -0.02). In the L2 pure condition, there was

minimal differences between words studied in repeated (logβ = -0.03) and varied (logβ = -0.07) sentences. When words were studied visually, there was a greater YES bias to words studied in varied sentences when words were presented in the L1 and L2 pure conditions (L1: repeated logβ = -0.71, varied logβ = -0.88; L2: repeated logβ = -0.34, varied logβ = -0.69). However, in the mixed-language study condition, a greater YES bias was observed for words studied in repeated sentences (logβ = -0.74) compared to varied sentences (logβ = -0.25).

Table 3.7: Mean (SE) signal detection scores in auditory and visually presented stimuli across language conditions in the semantic relatedness task

	d'				**Lambda Scores**			
	Visual		Auditory		Visual		Auditory	
	Repeated	Varied	Repeated	Varied	Repeated	Varied	Repeated	Varied
L1	1.61	1.63	1.18	1.02	0.58	0.56	0.44	0.42
	(0.12)	(0.12)	(0.12)	(0.12)	(0.09)	(0.09)	(0.73)	(0.78)
L2	1.33	1.41	0.72	0.72	0.49	0.47	0.31	0.38
	(0.13)	(0.12)	(0.13)	(0.12)	(0.09)	(0.08)	(0.78)	(0.67)
Mixed	1.42	1.40	0.80	0.77	0.55	0.53	0.31	0.41
	(0.13)	(0.14)	(0.13)	(0.14)	(0.09)	(0.10)	(0.66)	(0.80)

	Logβ			
	Visual		Auditory	
	Repeated	Varied	Repeated	Varied
L1	-0.634	-0.820	-0.192	-0.005
	(0.20)	(0.20)	(0.67)	(0.68)
L2	-0.419	-0.745	-0.078	-0.064
	(0.12)	(0.16)	(0.67)	(0.57)

Mixed	-0.737	-0.249	-0.156	0.018
	(0.20)	(0.12)	(0.54)	(0.66)

3.3 Discussion

Consistent with Experiment 1, an advantage in recall was observed for words that were studied in varied semantic contexts relative to a single, repeated context. Although previous research has examined auditory novel word learning in varying contexts (e.g. Nelson et al., 2005), this is the first study to provide direct evidence for the advantage of varied semantic contexts auditorily. To our knowledge, no previous study has directly compared the recall of words learned in varied or repeated sentences auditorily. The magnitude of the benefit for varied semantic contexts was similar in both L1 and L2.

As in Experiment 1, we predicted an advantage for studying novel words in mixed languages versus language pure environments. This prediction was not supported. Words studied in the L1 pure condition still had an advantage compared to words studied in mixed language environments. However, unlike Experiment 1, the effects of semantic variation persisted even in the mixed language study condition. Although participants in the sample were highly proficient bilinguals, they self-rated their proficiency in both languages higher in the auditory modality (i.e. speech comprehension: English $M = 9.17$, Spanish $M = 8.88$) than visual modality (i.e. reading comprehension: English $M = 8.31$, Spanish $M = 7.89$), $t(71) = -5.36$, $p < .001$ and $t(71) = -4.46$, $p < .001$, respectively. Words that were studied in both languages may have had slightly stronger activation of semantic features in L2 encounters, combined with stronger activation of traces to previous L1 exposures when those words were presented in a modality in which participants' proficiency was stronger in both languages. Thus, the activation

of traces to previous encounters in varied semantic contexts could spread more easily across languages.

The results of the auditory semantic relatedness task did not provide direct evidence for an advantage in words studied in varied semantic contexts. The fact that a time constraint was included provides converging evidence for our first interpretation of this null finding. That is, the high frequency synonyms included in the word-pairs provided a powerful retrieval cue. This retrieval cue may have activated a set of strong, similar features (for words learned in the same context repeatedly), or broader, perhaps weaker, features (for words learned in varied semantic contexts) equally as efficiently.

Unlike Experiment 1, there was no effect of semantic variation observed in the RTs of any of the language context conditions. However, there was a main effect of semantic contextual variation in the bias score data, in which a smaller degree of YES bias was observed when words were studied in varied semantic contexts than words studied in the same, repeated semantic context. Although the interaction was not significant, the reduction of response bias for words studied in varied semantic contexts was observed in the L1 pure, and mixed-language condition, but not in the L2 pure condition. This finding provides converging evidence that L1 traces were stronger in memory and more likely to be retrieved at test than L2 traces.

Because items and procedures were shared across Experiments 1 and 2, cross-experimental comparisons could be made on the differences between word learning in visual and auditory modalities. The results regarding the effects of language and semantic contexts are simply the averaged effects across Experiments 1 and 2. Thus, the advantage of studying words in L1 was replicated, as was the advantage of studying words in varied semantic contexts.

The novel contribution from the analyses performed across experiments is that a general advantage for visually presented stimuli was observed in both testing tasks. Participants generated a greater proportion of accurate responses in the meaning generation task for words that were learned visually compared to auditorily, and discrimination scores in the semantic relatedness task were greater in the visual modality. The results replicate previous findings demonstrating an advantage for visual stimuli in memory compared to auditory stimuli (Dean et al., 1988; McDermott et al., 2001), and specifically an advantage for learning new words visually (Nelson et al., 2005). One theoretical explanation for this finding is that orthography activates more strongly phonological codes during visual encounters with words than cross modal activation of orthographic codes during auditory word encounters (Borowsky et al., 1999). However, in the present study, the advantage for visually presented stimuli may also be due to the fact that there was no limit on the time or number of times participants could read or re-read the critical words in sentences. In the auditory experiment, participants only heard each critical word in isolation and in sentence contexts once each before moving on to the next trial.

Even so, the effect of modality did not appear to interact with the effect of language in which novel words were studied. In both visual and auditory experiments, words studied in L1 consistently elicited a higher proportion of correct responses in the meaning generation task. Furthermore, the magnitude of the advantage of studying words visually was consistent for words studied exclusively in L1, exclusively in L2, or in both languages. It appears that developing rich lexical-semantic representational networks for words depends in part on the extent that semantic features of a word are formed in learners’ dominant language regardless of whether they are encountered auditorily or in print.

In the semantic relatedness task, the analysis of bias scores revealed a three-way interaction between the modality, language, and semantic contexts in which words were encountered. Although there was an overall advantage for presenting words visually compared to presenting words auditorily, participants were less biased in their responses to auditory stimuli, especially when words were studied in varied semantic contexts. For words studied visually, there was a combined effect of language and semantic contexts such that the bias to respond YES was reduced when words were studied in both languages, and in varied semantic contexts. For words studied in the auditory experiment, there was a reduction of bias to respond YES when words were studied in varied semantic contexts, but only when at least some sentences were presented in L1 (in exclusively L1 and mixed conditions).

Chapter 4: General Discussion

The present study examined the impact of variability of language context on the learning of new words. The central hypothesis was that encountering new words across different languages enhances the memory trace due to added distinctive cues relative to encounters in just one language. If distinctiveness across encounters facilitates novel word learning, then encountering the same word in two languages (in two separate instances) should allow for resonant activation of semantic connections in both languages, thus providing learners with more potential connections to retrieve and/or generate meanings. A related hypothesis was that the distinctiveness of language context would differ across visual versus auditory modalities. Specifically the hypothesis was that the auditory modality would increase the distinctiveness of language cues since languages have unique phonemes and phonetic realizations. The central hypothesis was not supported. Across both experiments, performance was not superior for words presented in language mixed conditions relative to language pure conditions. Instead, performance was better for words presented in L1 pure conditions than L2 pure or mixed. This suggests that the key factor is the dominance of the language. That the mixed language condition was more beneficial than L2 may be due to the fact that at least some sentences were presented in the L1. Thus, the advantage of mixed language learning relative to L2 may be because of the increased proportion of studied items in the dominant language. Semantic information is more readily retrieved from L1, allowing for deeper encoding of semantic cues. Therefore, the best performance observed was for words studied in varying semantic contexts, but exclusively in L1. This afforded the learner multiple semantic cues that could be readily and more deeply encoded. Across both experiments, in both of the language pure learning conditions, an advantage for studying words in varied sentences was observed because the

salience of relevant semantic information is stronger. However, in the absence of a potentially rich semantic context, such as in the repeated sentence, the salience of language as an organizer in memory may increase. This was observed in the mixed language condition for Experiment 1. When participants read words in the same, repeated context in two languages, there was no significant difference with words read in varying sentences in the proportion of correctly generated meanings of those words.

The present study replicated the semantic context variation effect observed in single language word learning studies (Bolger et al., 2008; Balass, 2011). Namely, across both experiments, participants were aided by varying semantic contexts when encountering words exclusively in their L1, and exclusively in their L2. The magnitude of the advantage of encountering words in varied semantic contexts was greater in the meaning generation task relative to the recognition task. Having to generate or recall the meanings of recently learned words is likely to be more influenced by the availability of multiple memory chases relative to recognition. The recognition task did not require the same degree of retrieval of semantic features of the novel words, thus, responding could have been more easily based on the activation of a single episodic memory trace.

Additionally, the finding that an advantage of encountering words in varying semantic contexts was observed in bilinguals' L2 extends findings on studies investigating L2 novel word learning (e.g. Elgort et al., 2015). Although studies investigating L2 word learning in context logically extended the principles of the *Context Variability Hypothesis*, the present study is the first to directly compare the effects of encountering words in varying semantic contexts with encountering words in the same repeated context in the L2. The fact that the magnitude of the benefit of varying semantic contexts was similar in both languages is interesting in that it appears

that encountering a word in a variety of semantic contexts strengthens the links between semantic features within a word representation's network in both the L1 and L2.

It is important to note that the participants across both experiments were highly proficient bilinguals with already established lexicons in both languages. Previous research on building vocabulary in L2 learners has identified L2 lexical proficiency as a modulating factor in L2 word learning (Elgort et al., 2015; Elgort & Warren, 2014; Ferrel-Tekmen & Daloğlu, 2006; Horst et al., 1998; Pulido & Hambrick, 2008). Therefore, the effects of semantic variation may be attenuated in L2 relative to L1 in less balanced bilinguals.

It is possible that the influence of language context is the opposite of that hypothesized in the present study. Specifically, consistency in language may be more advantageous than variability. Other studies have demonstrated a benefit in consistency of non-semantic information, such as modality, across multiple encounters at study and test (Nelson et al., 2005). Perhaps consistency in the language context across encounters promotes more overlap between memory traces, facilitating semantic retrieval processes. If non-semantic information is consistent across encounters, activation of previous encounters and semantic connections to known words may occur faster or more efficiently. Thus, the effect of semantic context variability should occur when the surrounding language context is consistent across encounters. Indeed, this effect was observed in both visual and auditory word learning experiments.

Furthermore, the salience of non-semantic episodic information associated with previous encounters may have been weakened due to the spacing between each encounter (Francis, 1999; Glanzer & Duarte, 1971). Early research has demonstrated that as the space between repetitions at study decreases, the advantage in memory of words encountered in both languages at study increases (Glanzer & Duarte, 1971). In the present study, repetition of stimuli occurred after all

other words were presented in context. It might be the case that the non-semantic contextual environment (i.e. language membership) was not as relevant at retrieval when testing acquired knowledge of a word, because semantic information was more salient and actively searched for, whereas non-semantic information may have been weakly activated or even forgotten.

The present study does provide some preliminary evidence that varying the language in which words were encountered influences word learning. In experiment 1, there was no effect of semantic variation when words were studied in both languages. One interpretation of this finding may be an increase in the salience of non-semantic information (i.e. language context) as an organizer for lexical-semantic information in the absence of more semantically rich contexts. That is, bilinguals may rely more heavily on language contextual cues when there are fewer semantic contextual cues available at retrieval. Furthermore, studying words in both languages resulted in less biased responses in the semantic relatedness task. This was particularly true when the semantic context in which words were studied was also varied. Although no additive effects of varying language and semantic contexts across multiple encounters were observed in the accuracy and discrimination, varying the language in which a word is encountered did serve to reduce response bias above and beyond any effects of semantic variation. However, these findings must be interpreted cautiously, as they were only observed when words were studied and tested visually.

According to the *lexical quality hypothesis*, facilitation of a word representation at retrieval is enhanced when words have high lexical quality (Perfetti, 1985, 1992). Multiple encounters with words develops a single representation of those words. In order to be of high lexical quality, word representations must have consistent form to meaning mappings. Although the use of cognates was intended to increase overall quality of novel words, it may be the case

that encountering words in multiple language contexts had lower lexical quality because participants only had truly consistent form to meaning mappings half as often for words learned in both languages than words learned exclusively in one language. Furthermore, previous research has demonstrated that less-skilled learners dedicate more cognitive resources to lower-level comprehension processes (e.g. spelling or grammar), thus limiting the amount of resources dedicated to extracting meaning from context (Hart & Perfetti, 2008). Even highly proficient bilinguals operate more efficiently in one of their languages. If we consider bilingual L2 processing analogous to "less-skilled" learners, the pattern of results is consistent with the *lexical quality hypothesis*. During L2 presentations of words in context, participants may have dedicated precious cognitive resources to lower-level comprehension processes (such as decoding, spelling, or grammar), thus limiting the amount of resources dedicated to extracting meaning from context. Therefore, words learned exclusively in L1 were of higher lexical quality than words learned exclusively in L2; and words learned with some exposures in L1 were of higher quality than words with no exposures in L1.

The *Distributed Feature Model* (de Groot, 1992) allows for a similar interpretation of the findings. According to this model, access to semantic representations depends on the degree of overlap in lexical and semantic features in both languages. For previously unknown words encountered in L2 environments, it may have been less likely that semantic features were developed in the L1. Therefore, during testing, participants may have had fewer, and/or weaker, overlapping semantic features for words that were learned in the mixed language condition. This explanation is consistent with previous research demonstrating limited overlap of L1 and L2 semantic features in developing L2 vocabularies (Elgort et al., 2015; Finkbeiner, 2003; Jiang, 2000; Silverberg & Samuel, 2004; Frishkoff, Perfetti, & Westbury, 2009). Semantic knowledge

of words is developed incrementally. Thus, at a given time, a learner may have incomplete knowledge of meanings for some words (Frishkoff et al., 2009. In one study (Frishkoff et al., 2007), participants with strong vocabulary skills had higher accuracy on a lexical-decision task than less-skilled participants after studying extremely rare words in multiple contexts. This suggests that when learners have greater general vocabulary knowledge, the acquisition of semantic representations of newly acquired words is more robust. It might be the case that with limited exposure to words in L2 contexts, semantic connections were weaker within the L2, limiting the activation of related semantic knowledge in the L1.

A secondary goal of the present study was to compare how word learning across multiple encounters with words in different languages may differ as a function of the modality in which words were studied and tested. Previous research suggests that modal information (i.e. written or spoken) is encoded in memory traces (e.g. Nelson et al., 2005; Strobach, 2015) and that there is an overall advantage in memory for visually presented stimuli (Borowsky et al., 1999). The present study replicated previous research on learning words in sentences (Nelson et al., 2005), demonstrating an advantage for visual presentation of stimuli compared to auditory presentation.

Specifically, we hypothesized that presenting stimuli auditorily might activate language-specific phonological cues that might make the language context more salient, and more distinct across encounters with words. This hypothesis was not directly supported. However, the novel contribution of the cross-experimental comparison is that an advantage for visually presented stimuli was observed when words were studied exclusively in L1, exclusively in L2, and also when words were studied in both languages. This may be due, at least in part, to the auditory task being more difficult than the visual task. Specifically, participants in the auditory

experiment could not listen again to sentences at study the same way participants could re-read sentences.

The present study is the first to investigate word learning in bilingual sentence contexts. We have provided evidence that when bilinguals encounter words in either their L1 or L2 in a variety of semantic context, a greater degree of word knowledge is acquired. Although our predicted hypothesis that varying language contexts would result in greater semantic knowledge of a word due the spreading of semantic connections across languages was not supported, the language context in which a word is encountered does influence knowledge acquisition of a word separately (although perhaps not additively) from the degree of variation in semantic context. Perhaps non-semantic language features of the memory traces of word encounters are not as strong as semantic features, especially when the ultimate goal is learning the meanings of the novel words.

Although *instance-based* accounts of memory have been used throughout various sub-fields of cognitive psychology, they are not widely thought of in the bilingual and language learning literature. The present study informs current models of word learning in that various aspects of episodic memory are encoded in each trace of word encounters, and depending on the goal of the learning, different aspects of episodic or semantic information may be more or less salient. For example, when trying to learn the meanings of new words, semantic information included in the memory trace may be more relevant than linguistic or modal information. Future research may better elucidate the influence of non-semantic features of contextual word learning in two ways. First, an incidental word learning study may maximize the potential resonance activation of non-semantic components within the memory traces of previous encounters with words. Second, directly manipulating the spacing of repetitions of encounters with novel words

at study may provide further insight into the degree to which language membership is encoded in memory. If episodic features of word encounters are more salient when words are encountered more closely (in time) (i.e. Glanzer & Duarte, 1971; Francis, 1999), it might be the case that varying the language in which words are encountered will be more advantageous when words are encountered closer together in time.

References

Alameda, J. R. & Cuetos, F (1995). *Diccionario de frecuencias de las unidades linguisticas del castellano.* Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics.* Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M., (2008). Mixed effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

Balass, M. (2011). Learning words in context: an ERP investigation of word experience effects on familiarity and meaning acquisition (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh, PA.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97*, 548-564.

Bates, D. M. (2007). *Linear mixed model implementation in lme4*. Unpublished manuscript. Madison: University of Wisconsin.

Bates, D. M., & Sarkar, D. (2007). *Lme4: Linear mixed-effects models using S4 classes*, R package version 0.99875-6.

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.28, http://www.praat.org/.

Bolger, D. J., Balass, M., Landen, E., Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: an instance-based learning approach. *Discourse Processes, 45*, 122-159.

Borowsky, R., Owen, W. J., & Fonos, N. (1999). Reading speech and hearing print: Constraining models of visual word recognition by exploring connections with speech perception. *Canadian Journal of Experimental Psychology, 53*, 294-305.

Castles, A., Holmes, V. M., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *The Quarterly Journal of Experimental Psychology, 56*, 445-467.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25*, 1-18.

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods, 37*, 65-70.

Dean, R. S., Yekovich, F. R., & Gray, J. W. (1988). The effect of modality on long-term recognition memory. *Contemporary Educational Psychology, 13*, 102-115.

de Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1001-1018.

de la Riva Lopez, E. M., Francis, W. S. & Garcia, J. (2012). Repetition priming within and between languages in verb generation: Evidence for shared verb concepts. *Memory, 20*, 358-373.

Dufau, S., Grainger, J., & Holcomb, P. J. (2008). An ERP investigation of location invariance in masked repetition priming. *Cognitive, Affective, & Behavioral Neuroscience, 8*, 222-228.

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning, 61*, 367-413.

Elgort, I., & Piasecki, A. E. (2014). The effect of a bilingual learning mode on the establishment of lexical semantic representations in the L2. *Bilingualism: Language and Cognition, 17*, 572-588.

Elgort, I., Perfetti, C. A., Rickles, B., & Stafura, J. Z. (2015). Contextual learning of L2 word meanings: Second language proficiency modulates behavioural and event-related brain

potential (ERP) indicators of learning. *Language, Cognition And Neuroscience, 30*, 506-528.

Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning, 64*, 365-414.

Ferrell Tekmen, A., & Daloğlu, A. (2008). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals, 39*, 220-243.

Finkbeiner, M. S. (2003). Bilingual lexical memory: Towards a psycholinguistic model of adult L2 lexical acquisition, representation, and processing (Unpublished doctoral dissertation). University of Arizona, Tucson, AZ.

Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: semantic representation. *Psychological Bulletin, 125*, 193-222.

Francis, W. S., Fernandez, N. P., & Bjork, R. A. (2010). Conceptual and non-conceptual repetition priming in category exemplar generation: Evidence from bilinguals. *Memory, 18*, 787-798.

Francis, W. S., & Goldmann, L. L. (2011). Repetition priming within and between languages in semantic classification of concrete and abstract words. *Memory, 19*, 653-663.

Francis, W. S., & Strobach, E. N. (2013). The bilingual L2 advantage in recognition memory. *Psychonomic Bulletin & Review, 20*, 1296-1303.

Frishkoff, G. A., Perfetti, C. A., & Westbury, C. (2009). ERP measures of partial semantic knowledge: Left temporal indices of skill differences and lexical quality. *Biological Psychology, 80*, 130-147.

Fukkink, R. G., Blok, H., & de Glopper, K. (2001). Deriving word meaning from written context: A multicomponential skill. *Language Learning, 51*, 477-496.

Gallo, D. A., McDermott, K. B., Percer, J., & Roediger, H. L. (2001). Modality effects in false recall and false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 339-353.

Glanzer, M., & Duarte, A. (1971). Repetition between and within languages in free recall. *Journal of Verbal Learning and Verbal Behavior*, *10*, 625-630.

Hart, L., & Perfetti, C. A. (2008). Learning words in Zekkish: implications for understanding lexical representation. In E. L. Grigorenko, A. J. Naples (Eds.) , *Single-word reading: Behavioral and biological perspectives* (pp. 107-128). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Hintzman, D. L. (1984). MINVERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers, 16,* 96-101.

Horst, M., Cobb, T., Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*, 207-223.

Jenkins, J. R., Stein, M., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal, 21*, 767-787.

Jiang, N. (2000). Lexical development and representation in a second language. *Applied Linguistics, 21*, 47-77.

Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing: looking back and to the future. In J. F. Kroll, & A. M. B. de Groot (Eds.) , *Handbook of Bilingualism* (pp. 531-554). New York, NY: Oxford University Press, Inc.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492-527.

Lukatela, G. Frost, S. J., & Turvey, M. T. (1999). Identity priming in English is compromised by phonological ambiguity. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 775-790.

Lukatela, G. & Turvey, M. T. (1994). Visual lexical access is initially phonological: 2. Evidence from phonological priming by homophones and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 331-353.

Marsolek, C. (2004). Abstractionist versus exemplar-based theories of visual word priming: A subsystems resolution. *The Quarterly Journal of Experimental Psychology, 57,* 1233-1259.

Marsolek, C., Schacter, D. L., & Nicholas, C. D. (1996). Form-specific visual priming for new associations in the right cerebral hemisphere. *Memory & Cognition*, *24*, 539-556.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1155-1172.

McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*, 375-407.

McKeown, M. G. (1985). The acquisition of word meaning from the context by children of high and low ability. *Reading Research Quarterly*, *20*, 482-496.

McRae, K., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126,* 99-130.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238.

Morton, J. (1970). The role of memory in the attention process. *Acta Psychologica, 33*, 271-279.

Nagy, W., & Anderson, R. C (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*, 304-330.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*, 237-270.

Naish, P. (1980). The effects of graphemic and phonemic similarity between targets and masks in a backward visual masking paradigm. *Quarterly Journal of Experimental Psychology, 32*, 57-68).

Nelson, J. R., Balass, M., & Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Language and Literacy, 8*, 101-120.

Perfetti, C. A. (2007). Reading ability: lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357-383.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, R., Treiman (Eds.) , *Reading acquisition* (pp. 145-174). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Perfetti, C. A. (1985). *Reading ability*. New York, NY, US: Oxford University Press.

Perfetti, C. A., & Bell, L., (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 30*, 473-485.

Perfetti, C. A., Bell, L., & Delaney (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27,* 59-70.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. S. Garfien (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*, 67-86. Washington, DC; American Psychological Association.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C, Elbro, & P. Reitsma (Eds.) , *Precursors of functional literacy* (pp. 189-213). Amsterdam: John Benjamins.

Perfetti, C. A., Wlotko, E. W. & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1281-1292.

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review, 107,* 786-823.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

Pulido, D., & Hambrick, D. Z. (2008). The virtuous circle: Modeling individual differences in L2 reading and vocabulary development. *Reading in a Foreign Language, 20*, 164-190.

R development core team (2010). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing, http://www.R-project.org.

Reichle, E. D. & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading, 7*, 219-237.

Rodriguez-Fornells, A., Cunillera, T., Mestres-Misse, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3711-3735.

Rose, R. G., Rose, P. R., King, N., & Perez, A. (1975). Bilingual memory for related and unrelated sentences. *Journal of Experimental Psychology: Human Learning and Memory, 1*, 599-606.

Roth, S. F., & Perfetti, C. A. (1980). A framework for reading, language comprehension, and language disability. *Topics in Language Disorders, 1*, 15-27.

Seidenberg, M. & Tanenhaus, M. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 546-554.

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51*, 381-398.

Slowiaczek, L., Soltano, E., Wieting, S., & Bishop, K. (2003). An investigation of phonology and orthography in spoken-word recognition. *The Quarterly Journal of Experimental Psychology, 56*, 233-262.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99*, 3-21.

Strobach Oronoz, E. N. (2015). The effects of task demands and Word frequency on language source encoding (Unpublished doctoral dissertation). University of Texas at El Paso, El Paso, TX.

Taylor, R. S., & Francis, W. S. (2017). Between-language repetition priming in antonym generation: Evidence that translation-equivalent adjectives have shared conceptual representations across languages. *Memory, 25*, 344-349.

Unswort, S. & Pexman, P. (2003). The impact of reader skill on phonological processing in visual word recognition. *The Quarterly Journal of Experimental Psychology, 56*, 63-81.

Van Daalen-Kapteijns, M., Elshout-Mohr, M., & de Glopper, K. (2001). Deriving the meaning of unknown words from multiple contexts. *Language learning*, *51*, 145-181.

Van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*, 193-211.

Van Hell, J., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning, 47*, 507-546.

Van Orden, G. C (1987). A ROWS is a ROSE: Spelling, sound and reading. *Memory & Cognition, 15*, 181-198.

Van Orden, G., Johnston, J., & Hale, B (1988). Word identification in reading proceeds: From spelling to sound to meaning. *Journal of Experimental Psychology, 14*, 371-386.

Woodcock, R., W., Muñoz-Sandoval, A. F., Ruef, M. L., & Alvarado, C. G. (2005). *Woodcock-Muñoz Language Survey-Revised.* Itasca, IL: Riverside Publishing

Ziegler, J. C., & Ferrand (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bullietin Review, 5*, 683-689.

Vita

Justin Lauro earned his Bachelor of Science degree in Psychology from The University of Pittsburgh in 2012. He was subsequently accepted into the Language Acquisition and Bilingualism doctoral program in Psychology at the University of Texas at El Paso. While pursuing his doctoral degree, Lauro served as a research associate in laboratories in the Department of Psychology and the Department of Education. In 2014 he received his Master of Arts degree in Experimental Psychology. He also taught two different undergraduate courses in Psychology over five academic semesters. Lauro has presented his research at several conferences, and has received funding for travel through the Dodson Travel Grants from the College of Liberal Arts. He currently has one publication (citation provided below) and another manuscript under review.

Lauro, J., & Schwartz, A. I. (2017). Non-selective lexical access in sentence context: A meta-analytic review. *Journal of Memory and Language*, *92*, 217-233.

Permanent address: 509 Robin Hood Circle

Naples, Florida 34104

This dissertation was typed by Justin Gerald Lauro.