ABSTRACT

Title of Document: COMPARING SECOND LANGUAGE LEARNERS' SENSITIVITY TO ARABIC DERIVATIONAL AND INFLECTIONAL MORPHOLOGY AT THE LEXICAL AND SENTENCE LEVELS

Suzanne Freynik, Doctor of Philosophy, 2015

Directed By: Dr. Kira Gor, Professor, Second Language Acquisition

While L2 learners are less sensitive than native speakers to morphological structure in general (Clahsen & Felser, 2006; Jiang, 2007; Neubauer & Clahsen, 2009), researchers disagree about the roles different features of morphological systems play in determining the timecourse and accuracy of their acquisition by L2 learners. Some studies suggest that L2 learners process derivational morphemes in a more native-like manner than inflectional ones (Silva & Clahsen, 2008; Kirkici & Clahsen, 2013). Other research demonstrates accurate acquisition of L2 inflectional morphology as well (Gor & Jackson, 2013; Hopp, 2003; Jackson, 2008; Sagarra & Herschensohn, 2010). To date, few studies have directly compared L2 acquisition of inflectional and derivational morphology (Silva & Clahsen, 2008; Kirkici & Clahsen,

2013). Arabic verbs exhibit a system of derivational morphology whose function in constraining event structures and theta roles allows for comparably direct comparison with inflectional morphemes at the sentence level.

Forty-four L2 learners and thirty-three native speakers of Arabic participated in the current study, which used three behavioral tasks: a primed lexical decision task, an acceptability judgment task, and a self-paced reading task, to triangulate a picture of L1 and L2 Arabic learners' commands of derivational and inflectional morphology at the lexical and sentential levels. Results of the lexical decision and self-paced reading tasks indicated that both L2 learners and native speakers alike made use of Arabic derivational and inflectional morphological structure during lexical access and sentence processing. However, the acceptability judgment task found that L2 learners made far more accurate judgments about Arabic inflectional errors than about derivational errors. By contrast, native speakers made accurate judgments about both kinds of morphological errors. Thus, L2 learners' behavior regarding Arabic inflectional morphology was at least as native-like as their behavior regarding derivational morphology, if not more so, across tasks. This pattern of results accords with previous research that found accurate processing of inflectional morphology in proficient L2 learners. It also adds to a growing body of research suggesting that the distinction between derivational and inflectional morphology in Semitic languages may be more graded than it is in Indo-European languages (Boudelaa & Marslen-Wilson, 2000; Frost, Forster, Deutsch, 1997).

COMPARING SECOND LANGUAGE LEARNERS' SENSITIVITY TO ARABIC DERIVATIONAL AND INFLECTIONAL MORPHOLOGY AT THE LEXICAL AND SENTENCE LEVELS

by

Suzanne Freynik

Submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:

Professor Kira Gor, Chair
Professor Robert DeKeyser
Professor Jeff MacSwan
Professor Polly O'Rourke
Professor Colin Phillips

Dedication

For Barbara S.

I miss you.

Acknowledgements

It's no exaggeration to say that I wouldn't have even made it to the dissertation stage of my studies had it not been for my advisor, Dr. Kira Gor. As a researcher, she has always modeled an agility of mind and a depth of knowledge that I admire and can only aspire towards. As a mentor, she struck the perfect balance between patience and tough love, and her faith in me held me up when my own faith faltered.

I am also deeply indebted to Dr. Polly O'Rourke, who has been a mentor to me in both research and life for the past several years. I couldn't believe my good luck when I met her; in fact, this whole project started from a conversation with her. In everything but the official titles, she and Dr. Gor were both my co-advisors.

I'm grateful to Dr. Robert DeKeyser for being such a helpful, skeptical committee member and such a seriously kind person. It is no wonder he always has a line of students waiting to work with him. And I'm grateful to Dr. Colin Phillips for finding time for me, too, when he's not only one of the busiest people I've ever met but one of the most astonishing, inspiring intellects I've ever seen in action. Thanks also to Dr. Jeff MacSwan for agreeing at such short notice to be dean's representative, and for bringing a fresh and incisive perspective to questions that needed it.

Many thanks to all the other professors who also saw me through this process. Dr. Peter Glanville made time to talk to me, both about theory and his intuitions as a language instructor, and I always left our discussions feeling newly fascinated with Arabic morphology. Dr. Tonia Bleam shepherded me through my first qualifying paper years ago, and her warmth and curiosity and brilliance are still beacons to me. My gratitude also goes to Dr. Jeff Lidz. Really, the whole language science community at Maryland is full of people who I would have paid money just to listen to them thinking out loud. To have had the opportunity to argue and collaborate and eat lunch with some of them has been an unbelievable privilege.

Thanks to Fuad Saleh for his infinite patience as a tutor. Thanks, too, to everyone who helped me with these experiments, especially Chris Heffner, Alex Drummond, Nesrine Basheer, Yakov Kronrod, Danyal Najmi, Rawad Kalhaji, Nancy Crowell, Buthainah and Assma Althowaini, and to Amit Chavan who took a Javascript problem that had been flummoxing me for weeks and solved it so quickly. This work also wouldn't have been possible without Timothy Buckwalter's and Dilworth Parkinson's Frequency Dictionary of Arabic, or Dilworth Parkinson's Arabic language listserv.

I'm grateful to my peers and colleagues and classmates, too, especially Goretti Prieto Botana, Megan Gilmore, David Libber, Katya Solovyeva, Katie Nielson, Ilina

Stojanovska, Buthainah and Assma Althowaini, Stephen O'Connell, Peter Ostus, Eric Pelzl, Megan Masters, Alia Lancaster, Lana Cook, and Anna Chrabaszcz.

I wouldn't have survived the Maryland winters if it weren't for my DC family in the Takoma Park house, especially Katie Heffernan, Genevieve Fulco, Layne Garrett, and Novella Heffernan, and in the CritterDen house, especially Caitlin Teska, Ryan Carey, James Martin, Lindsay Eyth, David Combs, Ewan Dunbar, and Dru. You were patient with me when I was at loose ends, and you led me back to somewhere near sanity again.

I'm grateful from the bottom of my heart to Chris Menocal, who has been the voice of reason and the arms of comfort and the spirit of fun, who has shown me over and over again how to be strong and brave and kind, and who I could talk to forever about everything over coffee on weekend mornings. Also, he introduced me to Malcolm X Park on Sunday afternoons, and I owe him a whole world of thanks for that alone. (If you are ever sad in the summer in DC, just hold on until Sunday and then go to Malcolm X Park.)

Thanks to my Pennsylvania family for being so much sweeter and more supportive than some of those other families I heard about in the dissertation support group sessions I attended. I don't have words for how thankful I am to my Ma, who always encouraged me, who always wanted me to be happy, even when our definitions of "happy" were sometimes different. Thanks and love to Joe, Mike, Lou, and Jon. Thanks and love to Diane.

This part is tricky. I want to thank the ones I don't have categories for; you're some of the ones I need to thank the most, and there are too many of you to list, but I'll try. All the gratitude ever for Dayanna Moreno, Adrianna Moreno, Carolyn Keller, Linnea Minich, Evan Jones, Rachel Maddox, Leslie Quasius, Beth Sperber-Richie, Jessica Forsyth, Renee Garcia, Moses Sternstein, Barbara Schulz, Chrissie Faupel, Brian Segal, Ethan Sapperstein, Goretti Prieto Botana, Ewan Dunbar, Jamie Trowbridge, Jazmin Rumbaut, P.J. Rey, Erin and Aaron Pendergrass, Lauren Stansbury, Diana Rhodes, Tim George, Jose Gomariz, Geoffrey Wall, Stacy Warnick-Hesse, Lizette Baghdadi, Wayne Elise, Blue Bluekowski, Erica Madrid, Nekia Wright, Dave Salge, Evangeline Rosel, Farah Tebou, Laila, everybody on the FC Narwals, and all the Burners, every single Burner, and everyone who took the time to sit down with me and a coffee or a beer at some point and just remind me again why it's all worth it.

Table of Contents

List of Figures

List of Tables

1 Introduction

Research has shown that L2 learners are less sensitive than native speakers to morphological structure (Clahsen & Felser, 2006; Jiang, 2007; Neubauer & Clahsen, 2009). However, derivational morphology appears to be easier for L2 learners to acquire than inflectional morphology (Diependaele, Duñabeitia, Morris & Keuleers, 2011; Silva & Clahsen, 2008). The evidence for this discrepancy comes mainly from studies using Indo-European languages whose derivational morphologies are less productive than their inflectional morphologies, and may not even require decomposition to be processed (Clahsen, Felser, Neubauer, Sato & Silva, 2010). Arabic derivational morphology, by contrast, is rich and productive, and preliminary findings indicate that L2 learners of Arabic decompose words into their sublexical structures during lexical access (Freynik, Gor & O'Rourke, submitted), suggesting that neither productivity nor lack of obligatory decomposition can account for the relative facility L2 learners show in acquiring derivational morphology.

The goal of the current study is to determine whether second language learners are in fact more sensitive to Arabic derivational morphology than to Arabic inflectional morphology during lexical access, and whether that morphological sensitivity extends to sentence processing domains (as opposed to being limited to lexical processing). To this end, the current study examines L2 processing of Arabic derivational and inflectional morphology across three behavioral tasks: a primed lexical decision task, an acceptability judgment task, and a self-paced reading task. As Arabic verbal morphology allows for the comparison of derived and inflected variants of the same base forms, with sentence-level consequences for each, the current study is able to compare L1 and L2 learners'

sensitivity to both kinds of morphology at the lexical and sentence processing levels, in order to shed light on what aspects of the difference between derivational and inflectional morphology are most relevant for predicting their relative difficulties in L2 acquisition.

1.1 Overview

After discussing the experimental tasks typically used to examine morphological processing, the sections that follow will situate the current study by providing background on some of the general findings regarding morphological processing in native speakers at the lexical level (sections 2.1.2 and 2.1.3), and moving on to discuss the findings for L2 learners and how these differ (sections 2.1.4 and 2.1.5). Section 2.2 describes three theoretical approaches to L2 morphological processing, and the kinds of data each theoretical approach has been devised to account for. Section 2.3 turns to the literature on sentence-level morphological processing in L1 and L2 learners, focusing on representative results from studies using self-paced reading and eye-tracking methodologies. Following this review, it will be possible to outline the logic of the current study; after briefly explaining Semitic inflectional and derivational morphology (section 2.4), section 2.5 summarizes the findings in the literature regarding L1 Semitic morphological processing. Section 3 spells out three research questions, while section 4 explains a methodology for leveraging Arabic morphology to address them. Section 5 gives an overview of the results while section 6 contextualizes these in terms of the research questions and theoretical approaches laid out. Finally, section 7 suggests ways future studies can address the questions that the current study raises.

2 Review of the Literature

2.1 Morphology at the Lexical Level

2.1.1 Overview of Lexical Experimental Tasks

A central question guiding research into morphological processing has asked whether morphologically complex words are derived online via access to combinatorial rules (e.g., walk + ed = walked) or whether such words are retrieved full-form from the lexicon. Early investigations into this arena manipulated the base and surface frequencies of words in simple lexical decision tasks. A lexical decision task (LDT) is one in which a participant is presented with strings of letters and must decide, for each string, whether it represents a real word. The frequency of simple, monomorphemic words has been found to determine speed of lexical access in LDTs (e.g., participants recognize cat faster than sap because the former is a more frequently encountered word and is therefore more easily accessed).

For morphologically complex words, base frequency refers to how often a stem tends to appear, with any of its possible affixes (e.g., appearances of walk, walk-ed, walk-s, and walk-ing all count towards the base frequency of walk); whereas surface frequency refers to how often a specific surface form tends to appear (e.g., the surface frequency of walk-ed does not take occurrences of related forms like walk-ing into account). If morphologically complex words are retrieved full-form, the logic goes, the speed of access should be affected only by the frequency of that specific form (i.e., its surface frequency) whereas if they are derived by a rule that combines a stem with affixes, then the same stem is being accessed for every word that includes that stem and the speed of

access should be determined by a frequency measure that takes that into account (i.e., its base frequency, Pinker, 1999).

While proponents of different theories of morphological processing initially embraced LDTs with surface and base frequency manipulations as a diagnostic, a number of problems with this methodology were found to complicate interpretation of results. Base frequencies for irregularly inflected stems, for instance, could neither be satisfactorily calculated by excluding occurrences of irregular surface forms nor by including them (see Gor, 2010 for a detailed discussion of this problem). Furthermore, Baayen, Wurm and Aycock (2007) called into question the use of surface frequency effects as a diagnostic for full-form retrieval in the first place, presenting evidence that LDTs to low frequency, regular and morphologically complex words were affected by those forms' surface frequencies, even though such forms were in all likelihood being decomposed into their constituent morphemes during lexical access. In light of these difficulties with using frequency in simple LDTs as a diagnostic for morphological processing, another method is considered, namely primed lexical decision.

Lexical priming experiments rely on the logic that access to a given word will be facilitated following exposure to a related word. Thus, in priming tasks, related words (different degrees and types of relatedness are possible and are exploited in different experiments) are presented in sequence, with different amounts of time (and, in some paradigms, other stimuli) interspersed between them. The first word of a related pair is called the prime, and the second is the target. Three main kinds of primed LDTs have been shown to be specifically sensitive to effects of morphological relatedness. These are

delayed repetition priming, masked priming, and cross-modal priming. Examples will illustrate the use of each.

Delayed repetition priming is a technique that can be embedded in a seemingly simple LDT. Participants judge the lexicality of individually presented targets, however these items can be ordered in such a way that the prime and target pair occur in subsequent trials, or with a number of unrelated trials between them. Effects of semantic and phonological relatedness are apparent between adjacent or nearly adjacent prime-target pairs, but facilitation due to morphological relatedness is observable at greater distances; Drews & Zwitserlood (1995), for instance, found priming between morphologically related words (e.g., kellen 'ladles' primed kelle 'ladle') in Dutch and English, when 8 to 12 unrelated trials were interspersed between these. Evidence from Polish (Reid & Marslen-Wilson, 2000) & Hebrew (Bentin & Feldman, 1990) reinforces the same pattern; priming between morphologically related prime and target pairs can survive delays of up to 15 intervening trials, while effects of semantic and phonological overlap drop away. A criticism of this method, however, is that the delays between primes and targets may prompt participants to adopt strategic approaches (Monsel, 1985).

Masked priming was developed in part to avoid such strategy adoption by participants. In a typical masked priming paradigm, primes and targets are displayed in adjacent pairs, but prime words are displayed so briefly that participants are rarely conscious of having seen them. Nevertheless, preconscious lexical processing results in speeded RTs to targets following morphologically related primes.

Because masked priming is also susceptible to effects of orthographic overlap, it is important to include some baseline measure to control for this. Silva and Clahsen (2008), for instance, compared priming in morphologically related prime-target pairs (e.g., walked-walk) to an identity condition (e.g., walk-walk) as well as an unrelated baseline (e.g., pull-walk). They found that native English speakers showed the same degree of priming for morphologically related pairs as they did in the identity condition. Such results are often referred to as "full" morphological priming, in order to distinguish them from the "partial" priming that occurs when RTs in the test condition are faster than an unrelated baseline, but still significantly slower than the identity condition.

Another popular method for investigating morphological processing in lexical access is cross-modal priming. In a typical cross-modal priming task, the prime word is first presented auditorily, then the target word, to which the participant must make a lexical decision, is displayed visually. Because the prime and the target are presented in different modalities, it has been argued that cross-modal priming is especially suited to tap "central" lexical representations (Marslen-Wilson, Tyler, Waksler & Older, 1994; Marslen-Wilson, Ford, Older & Zhou, 1996). These are more abstract, modality-independent representations (as opposed to access representations, which include the features by which the lexical entry is identified and differ depending on the modality of access, graphemic or phonetic features, for instance). In other words, the idea behind cross-modal priming is to avoid using written primes for written targets because if reaction times are speeded in the experimental condition of such a task, it can be difficult to determine whether the priming arises from overlap between morphological representations or from simple form overlap at the orthographic level.

Through the use of each of these methodologies, research has begun to paint a picture of L1 morphological processing. The section that follows will outline this picture by touching on the more agreed-upon findings in the L1 literature. As research questions about L1 processing of morphology have been shaped by the kinds of morphology under examination, this discussion will first address findings for inflectional morphology and then move on to L1 research on derivational morphology.

2.1.2 L1 Processing of Inflectional Morphology at the Lexical Level

Inflectional morphology indicates grammatical information about a word, such as tense (e.g., -ed makes walk-ed past-tense) or plurality (e.g., -s makes pig-s plural). Crucially, it never changes the lexical class (e.g., it never makes a noun out of a verb) or fundamental meaning (i.e., dictionaries generally do not list inflected forms under separate entries). An early trend in the L1 literature suggested that processing of inflectional morphology was affected by regularity in languages like German, which exhibit clear distinctions between regular and irregular inflected forms. For example, in an investigation of the processing of regular and irregular German participles using cross-modal priming, Sonnenstuhl, Eisenbeiss and Clahsen (1999) found that regular inflection led to full priming in the test condition (e.g., ge-kauf-t ‘bought’ primed kauf-e ‘buy’ just as well as kauf-e primed itself). Irregular inflection, on the other hand, led to partial priming (e.g., ge-lauf-en ‘walked’ primed lauf-e ‘walk’ somewhat, compared to an unrelated baseline, but not as much as lauf-e primed itself). Similar results have been found for English past-tense verbs (Marslen-Wilson, Hare, & Older, 1993).

Researchers have explained these findings with reference to dual-mechanism theories of morphological processing (e.g., Baayen, Dijkstra, & Schreuder, 1997). The idea behind

such theories is that, while regularly inflected forms can be derived by way of combinatorial rules (e.g., walk + ed = walked), irregularly inflected forms must be stored as whole words (e.g., buy + ed ≠ bought, therefore bought, and its relation to buy, must be stored in the lexicon).

Such clean dissociation between regular and irregular forms, however, has recently been challenged by results from studies involving linguistic phenomena that exhibit more graded degrees of regularity. In a cross-modal priming experiment using inflected Russian verbs that differed in terms of regularity, morphological complexity, and type and token frequency, Gor and Cook (2010) found comparable priming for all verb classes. Neither the regularity nor the complexity of the inflectional paradigm significantly affected the degree of priming for native speakers. Similar lack of categorical distinction between regular and irregular forms has been found in Italian, another richly inflected language (Orsolini & Marslen-Wilson, 1997). Further still, Smolka, Zwisterlood and Rosler, (2007) have challenged Sonnenstuhl et al.'s (1999) initial findings for German verbal morphology, pointing out two major flaws in the study's design. First, Smolka et al. noted that the supposedly morphologically simple words used as targets (and as primes in the identity prime conditions) were in fact inflected forms, each consisting of the infinitive stem plus the –e suffix that signals first person singular present agreement. Thus, the identity priming condition did not truly represent a morphologically simple baseline. Secondly, word frequencies across conditions were not balanced, such that the frequencies of the irregular verbs were significantly higher than the frequencies of the regular verbs. This imbalance led to faster baseline (identity priming) RTs in the irregular condition, which Smolka et al. argue is

the main cause of the interaction Sonnenstuhl et al. observed between the verb types (regular vs. irregular) and the priming conditions (identity vs. inflected). Smolka et al. corrected for these shortcomings in a study of morphological priming in regular and irregular German participles and found that similar degrees of morphological priming could be observed for both regular and irregular participles (morphological priming also obtained between verbal stems and nonwords formed by illegal combinations of those stems with verbal affixes with which they do not actually occur in the lexicon). Smolka et al. concluded that these data constitute evidence of a single mechanism that processes both regular and irregular morphology via access to sublexical units. Such a progression from seemingly simple on towards increasingly nuanced findings is not exclusive to inflectional morphology; the research into derivational morphology is likewise full of controversy.

2.1.3 L1 Processing of Derivational Morphology at the Lexical Level

Whereas inflectional morphology generally signals a word's grammatical features, derivational morphology changes a word's meaning more fundamentally. The re- in re-heat, for instance, adds repetition to the original meaning of the verb. The –ness in dark-ness makes a noun out of an adjective. Research has shown that processing of derivational morphology in many languages is affected by semantic transparency. In a study of English derivational morphology using cross-modal priming, Marslen-Wilson, Tyler, Waksler and Older (1994) found that derived forms primed their stems so long as the meaning of that derived form could be understood in terms of its combined stem and affix. That is, a target-prime pair like involvement and involve gave rise to priming, whereas a pair like department and depart, where the relationship is not obvious, did not

prime. Similar patterns have been found for derived forms in Polish. In both cross-modal and delayed repetition priming paradigms, derived forms with transparent semantics primed their stems (e.g., bajk-o-pis-ar-stwo ‘fable-writing’ primed pis-a-c ‘to write’), while semantically opaque derived forms did not (e.g., jaloiec ‘juniper’ did not prime jalowy ‘futile’) (Reid & Marslen-Wilson, 2000).

Such effects of semantic transparency are not, however, universal. Recent research suggests they may be constrained by both experimental task and the typology of the target language(s). Whereas the abovementioned studies that used delayed repetition and cross-modal priming found clear-cut distinctions between semantically transparent and semantically opaque derivational morphology, masked priming studies tend to find subtler gradations in semantic transparency which correlate with similarly graded degrees of priming (see Rastle & Davis, 2008, for review). A recent and representative example is Diependaele et al.’s (2011) use of masked priming to test for facilitation at three levels of semantic relatedness: transparent (e.g., shipment-ship), opaque (e.g., department-depart), and unrelated-but-form-matched (e.g., freeze-free). Native speakers showed priming in both the morphologically related conditions (not in the form-matched condition), but the degree of priming was greater when the semantic relationship was transparent. Semantic transparency effects, however, have not been found in studies of priming in Semitic languages like Hebrew (Frost, Forster & Deutsch, 1997) and Arabic (Boudela & Marslen-Wilson, 2000). This discrepancy has been attributed to the relative productivity of derivational morphology in Semitic languages compared to Indo-European languages, and will be explored in greater detail below, in the discussion of research into Arabic morphological processing.

A number of studies have also pointed out that the processing of derivational morphology may rely on both combinatorial rule application and full-form storage. Clahsen, Sonnenstuhl and Blevins (2003) tested German native speaker's processing of -ung nominalizations in a cross-modal priming experiment, as well as in a simple LDT, along with -lien and -chen diminutives, and found that while derived forms tended to prime their stems (evidence of combinatorial rule application), the same derived forms showed surface frequency effects in the unprimed LDT (evidence of full-form storage). Neubauer and Clahsen (2010) found comparable results for their native speakers, this time using masked-priming (in combination with an unprimed LDT). The authors of both studies interpreted their results in light of a model of morphological processing wherein derived forms get separate lexical entries, which subsume their internal morphological structure. These "combinatorial entries" can give rise to surface frequency effects and morphological priming. In this way they account for the seemingly schizoid behavior of derived words (i.e., the fact that they exhibit effects of both storage and rule application). (This account is the basis for the Combinatorial Entries Hypothesis described in the introduction to the current study.)

Interestingly, there is also evidence that surface frequency affects processing of even regularly inflected forms above a certain frequency threshold (e.g., six per million words in English) in simple LDTs (Alegre & Gordon, 1999; Soveri, Lehtonen & Laine, 2007; see, however, Baayen, Wurm & Aycock, 2007, for evidence of surface frequency effects in very low frequency words as well, and an explanation of how such effects may obtain even when morphological decomposition is taking place), and that the type of nonwords used in the LDT may determine the kinds of frequency effects that emerge (Taft, 2004).

These findings lend support to the notion that full-form listings and combinatorial rules may not be at odds, and may in fact be processes that work in parallel, with the observable effects being the result of the faster process (Frauenfelder & Schreuder, 1992; Schreuder & Baayen, 1995). Which process turns out to be faster may depend on the frequency of the form, the productivity of the rule, and the demands of the task (e.g., identifying nonwords in one of Taft (2004)'s conditions required not just morphological decomposition but also recombination). In light of such findings, then, Clahsen, Sonnenstuhl and Blevins (2003)'s discovery of frequency effects in simple LDTs involving derived words does not necessarily indicate a qualitative difference between inflectional and derivational morphology. More substantial support for a qualitative difference between the two, however, comes from Bozic and Marslen-Wilson's (2010) review of fMRI evidence that the two kinds of morphology are processed in different areas of the brain. While inflectional morphology engages a left-lateralized decompositional subsystem, derivational morphology appears to be handled by a broader, bilateral network of brain regions. Bozic and Marslen-Wilson nevertheless point out that these conclusions are based heavily on research using English, which is a morphologically impoverished language in comparison to the richer systems of, for example, Slavic and Semitic. They note that additional research is necessary to confirm whether the observed patterns of neural activation would be born out in these languages as well.

If research into L1 morphological processing appears fraught with controversy, morphological processing in second languages (L2s) is even less well understood. The section that follows will lay out what findings have emerged thus far in the literature on

L2 morphological processing, addressing first inflectional morphology and then derivational, and noting ways in which these findings diverge from the literature on L1 morphological processing at the lexical level. Following this, the review turns to the literature on morphological processing at the sentence level.

2.1.4 L2 Processing of Inflectional Morphology at the Lexical Level

Most of the research so far into L2 morphological processing has focused on inflectional morphology. Sufficient evidence suggests that, while L2 learners do process inflectional morphology, their ability to make use of morphological information differs significantly from that of native speakers. In addition to being slower and less accurate than native speakers in LDTs, numerous studies have found morphological priming in L2 learners to be reduced or absent (Neubauer and Clahsen, 2009; Clahsen et al., 2010). For instance, Feldman, Kostic, Basnight-Brown, Durdevic and Pastizzo (2010) used both masked and cross-modal priming to examine native and L2 processing of regular and irregular past tense verbs. They found that while native English speakers showed priming in both regular and irregular priming conditions, L2 learners showed priming only for the regularly inflected verbs. In a similar vein, Basnight-Brown, Chen, Hua, Kostic and Feldman (2007) used cross-modal priming to look at processing of inflectional morphology, along a continuum of regularity, in L2 learners of English from Serbian and Chinese L1 backgrounds. While native English speakers showed facilitation for regularly inflected past tense primes (talked-talk), irregular nested stem primes (drawn-draw) and irregular stem-change primes (ran-run), no L2 learners showed priming for all three types of inflected forms. Chinese L1ers showed priming for regular past tense only, whereas Serbian L1ers showed priming for regular past tense and nested stem primes.

Comparable findings come from Clahsen, Sonnenstuhl and Blevins (2010), who cite evidence that English L2 learners from typologically distinct L1 backgrounds (German, Chinese and Japanese) displayed similar patterns in their difficulty with English morphosyntax. All learners were faster and more accurate at identifying errors involving case than errors involving agreement; however the degree of this discrepancy appears to vary greatly among L1 groups. It was greatest for Chinese learners, who were on average 527ms slower for items involving agreement and 17.2% less accurate. Japanese learners averaged 298ms slower and 9.7% less accurate, while German learners showed the smallest difference, averaging 162ms slower and 4.4% less accurate. As neither the raw means in accuracies or response times, nor the variance in the measures is reported, it is not possible to draw further conclusions from these apparent differences.

Clearer evidence that learners from different L1 backgrounds differ qualitatively in their processing of L2 morphology comes from Portin, Lehtonen, Harrer, Wande, Niemi and Laine (2008), who looked at two groups of L2 learners of Swedish. Native speakers of Hungarian had longer reaction times to medium and low-frequency inflected Swedish words (compared to monomorphemic controls), a slow-down that was interpreted as the cost of morphological processing. Native speakers of Chinese, however, had comparable reaction times to inflected and monomorphemic words alike at each frequency level, suggesting that they were accessing full-form listings of the inflected words. These differences in behavior between the two L1 groups are explainable in terms of transfer. Hungarian is an agglutinative language; like Swedish, it has rich inflectional morphology. The native speakers of Hungarian would be able to use similar routines to process the morphology in their L1 and their L2. Chinese, on the other hand, is an isolating language.

As it lacks inflectional morphology, its native speakers would be accustomed to accessing lexical items via full-form listings, and apt to transfer the same strategy to L2 lexical access as well.

Another factor that has proven relevant to L2 processing of inflected forms is morphological complexity. This factor was addressed by Gor and Cook (2010), in a study that used a real and nonce verb generation task, as well as a primed auditory LDT to examine the effects of morphological complexity and productivity in native speakers and L2 learners of Russian.

Russian is a language with rich inflectional morphology. Its verbs can be classified according to their stem endings, which in turn determine the ways that the stems change (or don't) when inflected with affixes. For instance, the underlying stem rabot-aj of the infinitive verb rabota-t, 'to work', undergoes automatic consonant truncation in forming that infinitive, but in the first person singular form, the suffix –u is added with no extra stem change, to give rabotaj-u. By contrast, the underlying stem ris-ova of the infinitive risova-t, 'to draw', undergoes suffix alternation to give the first person singular form risuj-u. Thus, –ova verbs like ris-ova are more morphologically complex than -aj verbs like rabot-aj. The –ova verbs are, however, morphologically unambiguous. Their inflected forms are totally predictable based on their stem ending. This morphological predictability contrasts with –aj verbs, which have the same infinitive ending as another class of verbs: those whose stems end in –a like pis-a, 'write'. Both pis-a and rabot-aj have infinitive forms that end in –at: rabota-t and pisa-t. An –at ending, then is not a clear indicator of which of the two paradigms a given verb belongs to. The declensional

paradigms differ, however, in frequency; the –aj paradigm is considerably more frequent than the –a paradigm.

Late L2 learners actually came closest to native-like accuracy in their generation of forms with unambiguous morphological cues, even when these forms were morphologically complex, such as those in the –ova class. This finding held for real verbs as well as nonce verbs, suggesting that learners could apply such rules online and not merely retrieve memorized forms. L2 learners' slower response times to verbs with complex allomorphy in the auditory LDT likewise suggested that they were engaging in online morphological processing.

2.1.5 L2 Processing of Derivational Morphology at the Lexical Level

While research examining L2 learners' processing of inflectional morphology has found reduced or absent effects compared to L1 controls (at times exacerbated by negative transfer and morphological complexity), research into L2 processing of derivational morphology has found comparably more native-like patterns. Silva and Clahsen (2008), for instance, compared processing of L2 inflectional and derivational morphology in the same study. They used masked priming to compare priming of -ed inflected verbs and –ness/-ity derived nominalizations in L2 learners from three different L1 backgrounds (German, Chinese and Japanese). Unlike native speakers, L2 learners showed no priming for inflected verbs, regardless of their L1 background. For derived nouns, L2 learners did show partial priming. Though this priming was smaller in magnitude than the repetition priming displayed by native speaker controls, it was nevertheless better than the total lack of priming L2 learners displayed when verbal targets were preceded by regular past-tense (-ed) inflected forms. Silva and Clahsen suggested that, because derivational morphology

does not require the same kind of combinatorial rule application that inflectional morphology requires, L2 learners are more apt to be able to process it. They claim that derivational morphology is stored differently than inflectional morphology in the mental lexicon; specifically, that derived forms are stored in "combinatorial entries" which can be retrieved full-form, but which also subsume internal sublexical structure. This claim will be revisited during the discussion of theoretical accounts of L2 morphological processing in section 2.2 below.

Additional evidence of L2 learners showing more native-like priming patterns for derivational morphology comes from a study (mentioned above in section 2.1.3) by Diependaele, Duñabeitia, Morris and Keuleers (2011). They used masked priming to look at processing of derivational morphology at three levels of semantic relatedness: transparent (e.g., shipment-ship), opaque (e.g., department-depart), and unrelated-but-form-matched (e.g., freeze-free). They found that L2 learners of English from three different L1 backgrounds (Dutch, French and German) showed the same pattern of priming as the L1ers in the same study. That is, they showed priming in both the morphologically related conditions (not in the form-matched condition), but the degree of priming was greater when the semantic relationship was transparent.

Further evidence of derivational morphological processing in L2 learners comes from Kim, Wang and Ko (2011). In an unmasked cross-language priming task, they found that L1 Korean learners of English showed facilitation of monomorphemic English target words when these were preceded by derived Korean words whose stems were translations of the English targets. This priming survived manipulations in lexicality and semantic interpretability, so priming was found not only for legal derived forms (e.g., the Korean

equivalent of attract-ive) but also for interpretable derived pseudowords (e.g., the Korean equivalent of attract-ivize) and even noninterpretable derived pseudowords (e.g., the Korean equivalent of attract-icide), whereas no priming was found for form-matched distractors with nonmorphemic endings (e.g., the Korean equivalent of attractive-el).

In summary, research into L2 morphological processing has generally found it to differ substantially from morphological processing by native speakers. Morphological priming in L2 learners tends to be reduced, by comparison, or absent altogether. That said, L2 learners tend to have more difficulty with inflectional morphology than with derivational morphology. Furthermore, L1 transfer can account for some observed patterns, to the extent that studies comparing morphological processing in L2 learners from different L1 backgrounds tend to find more native-like behavior among L2 learners whose L1s have similar morphology to the L2 in question. The section that follows discusses some theoretical explanations which have been proposed to explain these findings.

2.2 Theoretical Approaches to L2 Morphological Processing

The current study will focus on three theoretical accounts that may explain L2 learners' comparably less native-like behavior involving inflectional morphology (compared to derivational morphology): the Combinatorial Entries Hypothesis, the Uninterpretable Features Hypothesis, and the Sentence Level Dependencies Hypothesis. These three accounts begin with different claims put forth by Clahsen and Silva (2008), Jiang (2007), and Clahsen and Felser (2006), respectively, to explain observed L2 morphological deficiency. The original claims are extrapolated here to cover aspects of Arabic morphology, in order to generate predictions about the current study, as the following sections clarify.

It is worth underscoring that these accounts are explanations advanced by psycholinguists to explain a subset of morphological processing behaviors, and as such they abstract away from certain questions that face formal morphological theories. Formal morphological theories such as Minimalist Morphology (Wunderlich, 1995) and Distributed Morphology (Marantz, 1997; 2001) disagree about the exact representations of morphemes in the lexicon, the time courses of morphological processes and how these interact with syntactic processes. Distinguishing between formal theories at this level is beyond the scope of the current study. Within the context of the current study, the crucial distinction is between (de)compositional processes that entail access to sublexical structure versus full-form lookup processes that are blind to sublexical structure.

Crucially, the claim that both derivational and inflectional morphological processing involve accessing and representing sublexical structures is agnostic as to whether the two kinds of morphology are accessed and/or represented in qualitatively similar or different ways. Finer-tuned instruments are necessary to make such distinctions, and indeed, fMRI research suggests that derivational and inflectional morphological processing engage distinct brain regions (Bozic & Marslen-Wilson, 2010). It is left to future research to pin down theoretical specifics surrounding lexicon organization and the time course and manner in which morphological (de)composition interacts with syntactic processes. The current study is better framed in the context of the following three psycholinguistic hypotheses which focus on the sources of deficiency in L2 morphological processing.

The Combinatorial Entries Hypothesis describes an account first put forth by Silva and Clahsen (2008), who found that L2 learners of English exhibited priming between derived forms in a lexical decision task, when they showed none for inflected forms.

They argued that derivational morphology is more likely to exhibit priming in L2 behavioral tasks because derived forms get their own lexical entries in the lexicon, and such lexical entries are addressable with the declarative memory on which L2 learners rely. Silva and Clahsen refer to these entries as "combinatorial entries" because, although they can be retrieved full-form, they subsume the sublexical structure of a derived form. In this way they can account for the priming of a stem like dark after exposure to a derived form like dark-ness. As inflectional morphemes are stored in separate entries from the stems they modify, they cannot be retrieved as easily.

The Uninterpretable Features Hypothesis, by contrast, suggests that non-native-like L2 behavior involving inflectional morphemes arises not from the nature of the lexical entries that house them but rather from the features they encode. In a pair of studies that will be discussed in greater depth in Section 2.3.2 below, Jiang (2004; 2007) found L2 English learners' to be insensitive to errors involving plural –s. He explained these results in terms of feature interpretability; the L2 English learners in his studies were L1 speakers of Chinese, a language in which morphological plural marking is extremely rare. Note here that the distinction is between a language that marks a feature morphologically and a language that rarely does; the notion of plurality still exists in Chinese, but it tends to be expressed at the clause level.

Furthermore, Jiang pointed out that plural marking in English is often redundant in the sense that the information it encodes tends to be recoverable from other sources. L1 transfer and redundancy can work together to make a morpheme like –s "invisible" to L2 learners, resulting in selective integration, whereby certain L2 morphemes remain unacquirable.

A third possibility is that it is not specific unfamiliar features that lead to non-native-like morphological behavior in an L2, but rather sentence-level dependencies in general. This account will be referred to as the Sentence-Level Dependencies Hypothesis. This account is related to (but not as strong a claim as) Clahsen and Felser's (2006) Shallow Structures Hypothesis. The SSH maintains that L2 learners lack capacity for rule-based processing in the context of sublexical structures as well as syntactic structures. An alternate possibility that the current study will explore is that second language learners are able to engage in rule-based processing at the sublexical level, but that this ability breaks down at the (more complex, less constrained) syntactic level. Clahsen has suggested that the difficulty of processing sentence-level dependencies may correspond to the complexity of the structural relationship being signaled; for instance, he argues that L2 learners tend to make more accurate judgments about English pronominal case marking than about English subject-verb agreement because "SV agreement dependencies span the entire clause (and thus require comparatively complex structural scaffolding, whereas the objective case is assigned locally within the verb phrase" (Clahsen et al., 2010). This scenario would likewise predict more native-like L2 processing of derivational morphemes than of inflectional morphemes, as sentence-level dependencies are more often signaled by inflectional morphemes.

Of these three accounts, the latter two are difficult to test with tasks that tap primarily lexical-level processing. Interpretation of both morphological features and sentence level dependencies are better examined in the context of sentence processing tasks. The section that follows discusses inquiries into morphological processing at the sentence level.

2.3 Morphological Processing at the Sentence Level

2.3.1 L1 Processing of Inflectional Morphology at the Sentence Level

As O'Rourke and Van Petten (2011) explain, "Most of the world's languages use morphological agreement to flag relationships among words in sentences." Words are not processed in isolation in the wild, and the most ecologically valid view of morphological processing is afforded within a sentence-processing context. Morphology in sentential contexts is typically examined using a violation paradigm. That is, the morphological feature or relationship in question is isolated by way of an error; if researchers are interested in how speakers of a given language process number agreement, they test how those speakers respond to sentences with number agreement errors in them. While early studies (Johnson & Newport, 1989; Murphy, 1997; Whong-Barr & Schwartz, 2002) relied on offline measures like grammaticality judgment tasks, such measures provide only a broad, binary picture of morphological sensitivity, and are amenable to monitoring via conscious, metalinguistic strategies, a factor which is particularly relevant for L2 studies.

By contrast, self-paced reading tasks provide an online measure of processing difficulty in language comprehension. Typically, participants view a sentence that is masked by horizontal bars. By pressing a button, they reveal the words of the sentence, one word (or phrase) at a time, from left to right, allowing researchers to measure how long a participant spent reading each word. A number of factors may contribute to differences in reading times; L1 participants have been found to read more slowly immediately after encountering a semantically unexpected word (Vincenzi et al., 2003) or morphosyntactic error (Pearlmutter, Garnsey & Bock, 1999). For instance, Pearlmutter et al. found that

native English speakers tended to exhibit slower reading times in the region following an error in number agreement. That is, participants spent more time reading the word 'rusty' in an ungrammatical sentence like (1b) than in a grammatical sentence like (1a).

2.1a. The key to the cabinet was rusty from many years of disuse.

2.1b. *The key to the cabinet were rusty from many years of disuse

This slowdown relative to a matched, grammatical control condition demonstrates that participants are sensitive to morphological agreement during sentence comprehension. Further, the task in Pearlmutter et al. (1999) was ostensibly about sentence comprehension (participants were instructed to read for meaning, and half the questions were followed by Yes/No comprehension questions), which is to say, participants were not instructed to monitor sentences for grammatical errors. This suggests that native speakers access morphological information automatically, regardless of whether it is required by the task and regardless of whether they consciously attend to it.

2.3.2 L2 Processing of Inflectional Morphology at the Sentence Level

The first study to use self-paced reading to examine L2 learners' sensitivity to morphological agreement was Jiang (2004). Using a design similar to that of Pearlmutter et al. (1999), Jiang tested L2 learners of English whose L1 was Chinese to see how their reading times were affected by errors in plural marking and errors in verb subcategorization. The items testing for sensitivity to verb subcategorization errors involved contrasts like the one between (2a) and (2b) below.

2.2a. John encouraged me to go.

2.2b. *John supported me to go.

Jiang found that learners' reading times following errors involving plural marking were not statistically different from their reading times in the grammatical control condition. However, since learners' reading times were sensitive to verb subcategorization errors, L2 learners' insensitivity to the plural morpheme could not be said to arise from difficulty with the task itself.

In a follow up study, Jiang (2007) found L2 learners equally insensitive to plural -s marking when the morpheme was semantically incongruous (as opposed to the morpheme being the source of a grammatical agreement error), further clarifying that learners' difficulty was with the morpheme itself, and not agreement. Jiang explained these findings in terms of the learners' L1, Chinese, which generally does not express plurality using morphemes. He explained that L1 experience and morphological redundancy (syntactic and semantic) can work together to make a given L2 morpheme nonintegratable for learners from certain L1 backgrounds.

While Jiang explained L2 learners' difficulties with morphology in terms of L1 transfer, other self-paced reading studies highlight the role L2 proficiency plays in predicting the native-likeness of morphological processing. Hopp (2006) presented L2 learners of German whose L1 was English with sentences involving relative clauses with either subject-first or object-first word order. Telling the two constructions apart required correctly interpreting the case-marking on the nouns. Hopp found that while all the L2 learners in the study were native-like in their speeded judgments of the sentences, only L2 learners with near-native proficiency responded to the (pragmatically marked) object-first word order with slower reading times in the region following the case-marked noun. L2 learners whose proficiency was only at the advanced level, by contrast, showed

slower reading times at the last word of the sentence in the object-first condition only, suggesting that they may have waited until the end of the sentence to attempt a reanalysis. Jackson (2008) replicated these findings using wh- questions wherein both object-first and subject-first word orders were pragmatically unmarked options. Instead, the slower reading times in the object-first condition of Jackson's study resulted from a garden path effect because the first nouns always had ambiguous case-marking, and it is nevertheless the case that subject-first word order is preferred (if optional) in German.

Hopp's (2006) and Jackson's (2008) evidence that English L1, German L2 learners are sensitive to German case-marking would seem to contradict Jiang's contention that L2 morphemes tend to be unacquirable when they correspond to features which are not morphologically realized in the L1. One possible explanation for this discrepancy may be that English pronouns exhibit case-marking even though English nouns do not, such that case features might not be as inaccessible to L1 English learners as number features are for L1 Chinese learners. A second difference between the studies is that in Hopp's and Jackson's studies, case-marking information was necessary to interpret the target sentences, whereas plural marking was not necessary to interpret the sentences in Jiang (2004b) and (2007). In this sense, the plural morpheme examined was more redundant than the case-marking morphemes. A third possibility is that the learners in Hopp's and Jackson's studies may have simply been more proficient than the learners in Jiang's (2004) and (2007) studies. Only Hopp's near-native participants showed slow-downs in the spillover region; participants who were merely advanced showed slow-downs at the end of the sentence. As Jiang did not report RTs for sentence-final words, it is unclear how his L2 participants might have compared in this respect.

Additional evidence for the role of L2 proficiency comes from Sagarra and Herschensohn (2010), who found that while all the L1 English learners of Spanish who participated in their study could accurately identify gender and number errors in agreement in an offline acceptability judgment task (AJT), only the higher proficiency group (in this case, intermediates) demonstrated sensitivity to these errors in the self-paced reading task. Further, gender agreement is a linguistic phenomenon that is not present in these learners' L1, English, suggesting that lack of feature familiarity from the L1 may not constitute the same disadvantage in all L1-L2 pairings. Like case-marking, however, English does exhibit gender marking on its pronouns, so while morphological gender agreement is not present in English, one cannot say the feature is wholly unmarked in the language. The effect of different degrees of similarity in feature marking between the L1 and the L2 is still largely an open question, however, the evidence reviewed above suggests that familiarity-from-the-L1 and proficiency level interact to determine the native-likeness of L2 morphological behavior when it comes to inflectional morphology in sentence contexts.

Fewer studies have investigated the factors that affect derivational morphological processing in sentence contexts, and most of the studies that have done so have relied on base and surface frequency effects as diagnostics of morphological processing. One recent study uses a violation paradigm like those described above to examine derivational morphological processing. All of them are L1 studies. This research is reviewed in the section that follows

2.3.3 L1 Processing of Derivational Morphology at the Sentence Level

One of the earlier comparisons of derivational and inflectional morphological processing in sentence contexts comes from a self-paced reading study by Randall and Marslen-Wilson (1998). They compared residual reading times for words with morphological structure to residual reading times for monomorphemic English words, in both high and low surface frequency conditions across two experiments: one for derived words and one for inflected verbs. Randall and Marslen-Wilson found that words with morphological structure tended to have longer residual reading times than did monomorphemic controls, and that morphological structure contributed to reading time, independent of surface frequency. In a third experiment, they tested the effect of constraining context on reading times for novel words derived using affixes at varying levels of productivity and found that, while novel words led to longer reading times in less constrained contexts, contextual constraint shortened the reading times for novel words with productive affixes. For example, the novel derived word 'listy' is read more quickly in sentence 2.3b below, because its context more specifically anticipates 'listiness' than the context in sentence 2.4b below, in which it is read more slowly.

Strong pragmatic and syntactic constraint:

2.3a. John’s speech to the conference was filled with point after point.

2.3b. He began some tedious and LISTY/WORDY demands for better working conditions.

Weak pragmatic and syntactic constraint:

2.4a. John decided it was time to make the strength of his feelings clear.

2.4b. He began some LISTY/WORDY demands for better working conditions.

Randall and Marslen-Wilson interpret these findings as evidence against a strictly modular model of sentence processing in which lexical, syntactic, and semantic processes proceed sequentially and do not interact. That sentential context facilitates the recognition of novel, morphologically complex forms is taken by Randall and Marslen-Wilson as evidence that the different linguistic levels of processing do interact during reading.

Another comparison of derived and inflected forms in sentence contexts looked at the relative contributions of base and surface frequencies to reading times for both kinds of morphology. Using an eye-tracking methodology, Niswander, Pollatsek & Rayner (2000) manipulated the base and surface frequencies of derived and inflected words embedded in sentences and found that, while base frequency predicted gaze durations for derived words, gaze durations for inflected words were most reliably predicted by surface frequency. Among the inflected words, base frequency contributed to gaze durations only for the inflected nouns, not the inflected verbs. Niswander et al. speculated that this difference between the inflected nouns and inflected verbs might have come from the most common associations of those inflected forms' morphological stems. In the inflected noun condition, the stem was almost always still a noun. In the inflected verb condition, however, many of the verbs had stems that appeared more commonly as nouns; for instance, ‘handed’ is a synonym for 'passed' but 'hand-' is most frequently interpreted as a body part, not an action.

To explain the greater role base frequency played in predicting gaze durations for derived words than for inflected words, Niswander et al. appealed to affix length as a mitigating factor. Specifically, English derivational affixes tend to be longer than English inflectional affixes. The authors argued that this makes them more salient and likely to be

noticed early in the time course of lexical access. Additionally, initial fixations (i.e., the first place where the gaze lands when reading a given word) on longer words tend to fall further from the left edge and thereby closer to the morpheme boundary. Niswander et al. suggest that their results are best accommodated by a dual-route, race model of morphological processing, in which morphological decomposition and full form lookup proceed in parallel for any given morphologically complex form. The frequency effects observed, then, point to the route that won the race; base frequency effects indicate that morphological decomposition was faster, while surface frequency effects indicate that full form lookup was the faster route. In Niswander et al. (2000), factors such as affix length and stem homonymy contributed to the relative speed of the morphological decomposition route, making it less efficient than full form lookup in the inflected verb condition.

Additional support for this interpretation came from a follow up study; Niswander-Klement and Pollatsek (2006) manipulated the lengths, as well as the base and surface frequencies, of various derived English words and confirmed that, for longer words, base frequency was more predictive of gaze duration. Meanwhile, surface frequency was more predictive of gaze duration for shorter words. Similar findings are attested in Dutch; Kuperman, Bertram & Baayen (2010) found that gaze durations for derived words with suffixes tend to be best predicted by base frequency when the suffix is long, and by surface frequency when the suffix is short. Furthermore, Kuperman et al. also found that gaze durations were longer for words wherein a relatively productive stem was combined with a relatively nonproductive affix or vice-versa. Gaze durations were shorter when a word's stem and affix were comparably productive. This latter finding is hard to explain,

and the authors suggested it may have to do with the time course at which information from the different sublexical units becomes available. That is, it may be ideal for the decomposition route when information for both the stem and affix become available at roughly the same time. However, the authors note that more research is needed to get to the bottom of this effect.

The above examples demonstrate that using base and surface frequency as diagnostics for morphological decomposition and full form lookup respectively is a complicated enterprise, since factors like relative productivity, word length, suffix length, stem homonymy and contextual constraint all seem to play mitigating roles in determining which type of frequency effect is more strongly attested. Additionally, as noted in Section 2.1, there is some disagreement among researchers as to exactly how even straightforward frequency effects should be interpreted.

A further difficulty with base/surface frequency diagnostics, concerns the fact that L2 learners, particularly in the classroom, tend to encounter L2 lexical items in distributions that differ from the distributions in which those lexical items occur in more naturalistic settings. Thus, if L2 learners do not demonstrate the same kind of frequency effect that L1 learners show, it can be difficult to determine whether the source of the discrepancy is a different kind of processing route or mechanism in the L2 learner, or whether it is simply that the L2 learner's experience comes with its own (different) set of frequency counts. For all of these reasons, it is worth returning to the violation paradigm methodology for examining derivational morphological processing.

The only study to date that examines derivational morphology in sentential contexts using a violation paradigm comparable to the one employed in Pearlmutter et al. (1999) comes from Clahsen and Ikemoto (2012). In it, they compare Japanese deadjectival nominalizations formed with the suffixes -mi and -sa, in sentence contexts and in isolation. Both forms involve straightforward concatenation with no allomorphy. Crucially, while -sa nominalizations are productive and have strictly compositional meanings, -mi is a much less productive affix, and the nouns it derives in have specific and often idiosyncratic meanings. For example, atataka-i is the adjective for 'warm'. Atataka-sa, then, denotes simply 'the state of being warm', whereas atataka-mi carries the idiosyncratic meaning of 'warmth' as a personality trait.

Both -mi and -sa nominalizations exhibited surface frequency effects in an unprimed LDT, and both benefited from stem priming in an LDT with masked priming. Clahsen & Ikemoto took these results as evidence that -mi and -sa forms are processed similarly at the lexical level, via access to combinatorial entries. Where the forms differed was at the sentence level. Clahsen and Ikemoto used eye-tracking to examine reading times for both -mi and -sa nominals when these were embedded in sentences that either specifically licensed the -mi form's idiosyncratic meaning (such as in sentence 2.5a below), or sentences that did not (such as in sentence 2.5b).

2.5a. Kokyu wain-wa kutiatari-ga yawaraka-i. Daremo-ga sono yawaraka-mi-o mitomemasu.

‘Vintage wine has a smooth taste. Everyone approves of this smoothness.’

2.5b. Kokyu umoubuton-wa yahari yawaraka-i. Daremo-ga sono yawaraka-sa-o mitomemasu.

‘It has to be said that the luxury duvet is soft. Everyone approves of this softness.’

Gaze durations on -sa nominalizations were similar across sentence contexts, whereas gaze durations on -mi nominalizations were longer in the non-mi-licensing context condition. (These longer reading times corresponded to the lower acceptability ratings native speakers gave to sentences in which -mi nominalizations were embedded in non-mi-licensing contexts; -sa nominalizations were judged acceptable in both kinds of context sentences.) The authors argue that these results demonstrate that, while the representations for both kinds of Japanese nominalizations are similar at the lexical level, they differ at the lemma level where semantic properties are specified. Most crucially, however, this study demonstrates that semantic anomalies that arise due to inappropriate derivational morphology in sentential contexts will lead to slower reading times in native speakers, just as errors in inflectional morphology do.

As this review of sentence-level research into morphological processing demonstrates, there is comparably little work examining the processing of derivational morphology in sentential contexts anywhere, and none in the L2 domain. This is in part because it is difficult in many languages to set up a violation paradigm using derivational morphology that results in the same kind of sentence-level anomalies that result from faulty agreement between inflectional morphemes.

As Section 3 on the logic of the current study will elaborate, Arabic sublexical structure allows for the opportunity to examine derivational and inflectional morphology on more comparable footing. Before explaining exactly how this comparison will be laid out,

however, it will be necessary to describe how Arabic derivational and inflectional morphemes work.

2.4 Arabic Verbal Morphology

2.4.1 Arabic Verbal Inflection

Arabic verbal inflections are generally affixed to verb stems as prefixes and suffixes in a manner similar to the concatenative morphology of Indo-European languages. Arabic verbs agree with their subjects in person, number and gender; Table 2.1 depicts the affixes appropriate to ten Arabic pronouns.

Table 2.1 Arabic Verbal Agreement

Person	Number	Gender	Imperfect Tense	Perfect Tense
1st	singular	neutral	**a**fʕalu	faʕal**tu**
	plural	neutral	**na**fʕalu	faʕal**naa**
2nd	singular	masculine	**ta**fʕalu	faʕal**ta**
		feminine	**ta**fʕal**iina**	faʕal**ti**
	plural	masculine	**ta**faʕal**uuna**	faʕal**tum**
		feminine	**ta**fʕal**na**	faʕal**tunna**
ʕrd	singular	masculine	**ya**fʕalu	faʕal**a**
		feminine	**ta**fʕalu	faʕal**at**
	plural	masculine	**ya**fʕal**uuna**	faʕal**uu**
		feminine	**ya**fʕal**na**	faʕal**na**

2.4.2 Arabic Derivational Morphology

In contrast to Arabic's mostly concatenative verbal inflectional morphology, the derivational morphology of Arabic is templatic. This means that words are composed of at least two morphemes: a root and a pattern (also called a template). Roots are made up of consonants (usually three) and carry a word's semantic gist. Patterns are composed mainly of vowels (though they may also include some consonants) and provide both

phonological structure for the word and syntactic information about its role in a sentence. Words are derived by interleaving a root with a pattern; patterns include slots into which the consonants of the root fit when the two are combined. For example, the word ʕaalim, 'scholar', is derived from the root 'ʕ-l-m' and the pattern faaʕil (traditionally, patterns are written by substituting the three consonants 'F-ʕ-L' where the three consonants of the triliteral root would go). The root ʕ-l-m indicates the semantic field of knowledge, learning, and information. The pattern faaʕil is the pattern for active participles.

By combining the same root with a pattern for active verbs (faʕala) you get ʕalama, 'he knew'. If you combine it with a pattern for causative verbs (faʕʕala), you get ʕallama, 'he taught' or 'he informed'. If you combine it with a pattern for adjectives (faʕeel), you get ʕaleem, 'informed' or 'scholarly', and if you combine it with a pattern for passive participles (mafʕuula), you get maʕluuma, 'fact' or 'that which is known'. The matrix in Table 2.2 further illustrates this system of derivation.

Table 2.2 Arabic words derived from roots (columns) and patterns (rows)

	ك ت ب k t b related to writing	ع ل م ʕ l m related to learning	د خ ل d x l related to entering
فعل faʕala (_a_a_a) past tense verb	كتب kataba he wrote	علم ʕalama he learned	دخل daxala he entered
فاعل faaʕil (_aa_i_) active particple, doer of verb	كاتب kaatib writer	عالم ʕaalim scholar	داخل daaxil intrinsic part
مفعل mafʕal (ma_ _a_) place verb happens	مكتب maktab desk	معلم maʕlam marker, signpost	مدخل madxal entrance, foyer

The apparent systematicity of this matrix is, however, somewhat misleading. Like derived forms in other languages, the compositional semantics of Arabic words are not always synchronically obvious. That is to say, while nearly all Arabic content words can be decomposed into a root and a pattern, the meaning of a given word is not always interpretable as the sum of those morphemes. The root gh-r-b, when combined with different patterns, gives rise to the words sunset (maghrib), strange (ghareeb), and exile (gharba). Imagining a diachronic accumulation of meanings associated with going away, sunset, the west, foreigners, and oddness is an interesting thought exercise, but that such disparate associations should have psychological reality for modern speakers is far from given. Psycholinguistic research in this arena, however, as Section 2.5 below will explain, suggests that speakers do access the root as a separate morpheme, even when its contribution is semantically opaque, a feature which distinguishes Semitic languages from Indo-European ones. The focus of the current study is not on these idiosyncratic/ historical associations, however, but on a subset of derivational forms whose contributions to lexical meaning are often (if not always) more predictable. It will be argued that their relative systematicity makes the ten Arabic verbal patterns more comparable to inflectional morphology than many other kinds of derivational morphology.

2.4.3 Arabic Verbal Patterns: Ten Forms

The previous section described how Arabic words are formed by interleaving root and pattern morphemes. Pattern morphemes carry phonological and syntactic information like word class. A specific subset of pattern morphemes comprises the ten verbal forms that specify the argument and event structure for a given verb.

Form I, faʕala, is often described as the “basic” or “general” meaning of a given root in verb form. Some examples of Form 1 verbs are xaraja, ‘to leave’, ʕamala, ‘to do/make/work’, qaTaʕa, ‘to cut’, and jamaʕa, ‘to gather’. Form II, faʕʕala, has a causative and sometimes intensive meaning, such that xarraja, a causative derivation of ‘to leave’ means ‘to graduate (someone)’. Form III, faaʕala, has an associative meaning, such that ʕaamala, the associative derivation of ‘to do / to work’ means ‘to deal with’. Form IV, afʕala, like Form II, tends to have a causative meaning; for example, axraja means ‘to expell (someone)’. And while saqaTa (Form I) means ‘to fall’, asqaTa (Form IV) means ‘to drop (something)’. Form V, tafaʕʕala, generates a reflexive meaning and has an intransitive argument structure. The Form V derivation of ‘to gather’ is tajammaʕa, ‘to congregate together’. Form VI, tafaaʕala, has a reciprocal meaning; the Form IV derivation of ‘to work’ is taʕaamala, ‘to deal with each other’, and the verb that means ‘to exchange’, tabaadala, is a Form VI verb. Form VII, infaʕala, has an anticausative meaning, which is similar to passivization except that no external actor is implied. The anticausative derivation of ‘to break (something)’ is inkasara, ‘to become broken’, and the anticausative derivation of ‘to open (something)’ is infataHa, ‘to become open’. Form VIII, iftaʕala, is also an anticausative form, but it describes something undergoing an internally-caused process, in contrast to Form VII which tends to describe something instantaneously changing states (e.g., breaking, opening). The Form VIII derivation of ‘to spread (something)’, is intashara, ‘to spread (by itself)’, while the Form VII derivation of ‘to burn (something)’, is ihtaraqa, to burn (by itself). Form IX, ifʕalla, is very rare and has to do with acquiring an attribute, almost always a color; iHmarra, for instance, is the Form IX derivation of the root for ‘red’ and means ‘to turn red’. Form X,

istafʕala, has a considerative or requestive meaning. While kashafa (Form I) means ‘to reveal or unveil’, istakshafa (Form X) means ‘to explore’, and while hawiya (Form I) means ‘to love’, istahwaa (Form X) means ‘to seduce or enchant’. Table 2.3 summarizes these forms.

Table 2.3 Ten Verb Form Patterns

	Form	Meaning
1	faʕala	basic
2	faʕʕala	causative / intensive
3	faaʕala	associative
4	afʕala	causative
5	tafaʕʕala	reflexive
6	tafaaʕala	reciprocal
7	infaʕala	anticausative (process)
8	iftaʕala	anticausative (state change)
9	ifʕalla	attributive (colors)
10	istafʕala	requestive /considerative

This is a simplified description of the ten verb forms, as they are typically presented in Arabic textbooks. While it captures much of the systematicity apparent among verbs derived from the same root, the picture is actually more complicated. A subtler and more thorough discussion appears in Glanville (2012). Some key insights from his account include the observation that the ten forms do not actually specify semantic radicals like causativity or reflexivity so much as they specify the shape of an event structure. Thus, what Form IV does is specify an external actor; causative meaning is the result of

combining an external actor with certain kinds of events. Similarly, the roots do not contribute fixed meanings that plug into the verb forms. Rather roots designate broader semantic spaces, which are constrained by the requirements imposed by a given verbal form. For example, the root S-w-r has to do with images and pictures. Plugging this root into the Form II pattern, the result is Sawwara, to photograph, but the same root in Form V yields taSawwara, 'to imagine'. If Form V were simply a reflexivization of Form II, as some textbooks suggest, taSawwara should mean 'to photograph oneself'. Thus, taSawwara is better understood as the intersection of the semantic space of things related to imagery, and an event structure that designates an actor who is affected by his own action.

Motivating one analysis of Arabic verbal patterns over another is beyond the scope of the current study. A more important question at present concerns whether these morphemes have psychological reality. As Section 2.5 will describe, the evidence so far suggests that they do.

2.5 Psycholinguistic Perspectives on Arabic Morphology

2.5.1 L1 Processing of Arabic Derivational Morphology at the Lexical Level

Some of the first psycholinguistic evidence for the distinct representations of Arabic roots and patterns comes from Prunet, Beland and Idrissi's (2000) case study of an Arabic/French bilingual patient with aphasia, called ZT. ZT completed word reading, picture-naming, spoken repetition tasks in both his languages, and the authors found that he produced far more metathesis errors in Arabic than French. Further, these errors consisted almost exclusively of permuting root consonants; they rarely affected the

patterns (whereas the few metatheses he produced in French affected vowels and consonants indiscriminately). The authors concluded that his behavior was evidence that Arabic root consonants "float" at some level of representation in the minds of native speakers, and drew supporting connections with the observed permutability of Arabic root letters in tongue-slips among neurotypical native speakers, as well as in Arabic word games.

Further evidence for the special status of Arabic roots comes from Perea, Mallouh and Carreiras's (2010) investigation of transposed-letter (TL) priming in Arabic. The background for this experiment is a body of findings for Indo-European languages like English and Spanish, in which nonwords created by transposing two medial letters in a real word will prime that real word (e.g., jugde primes judge; Perea and Lupker, 2003). The authors found that Arabic prime words transposing the letters of the target words will speed RTs to those targets only when the transposition affected the order of the pattern letters. Transpositions affecting root letter order did not. The authors explained these findings in terms of the important role that roots play in Arabic lexical access. However, caution is appropriate in comparing their findings to those from the Indo-European studies, because while the latter employed nonword primes for real word targets, Perea et al. used all real word primes for real word targets. Thus, the transposed pattern-letter condition was also a morphological (root) priming condition, whereas the transposed root-letter condition was not.

Much of the current knowledge about Arabic morphological processing in healthy adult native speakers comes from a series of studies by Boudelaa and Marslen-Wilson (2000; 2001; 2004; 2005; 2011). Through approximately ten years of lexical priming research,

they found root priming to be the fastest and most robust morphological effect in native speakers of Arabic. Furthermore, they found root priming not to be constrained by semantic transparency, and to obtain in spite of allomorphic variation.

The speed of root priming was established in a masked priming experiment wherein the stimulus onset asynchrony (SOA, or the time between the moment when the prime word flashes on the screen and the moment it is replaced by the target word) was manipulated between subjects (Boudelaa and Marslen-Wilson, 2005). The four SOA conditions were 32ms, 48ms, 64ms and 80ms. Root priming was evident at the shortest SOA, whereas pattern priming did not emerge until 48ms. Root priming was also evident in all the SOA conditions, whereas pattern priming was fleeting (evident at only 48ms in the verbal condition, at 48 and 64ms for nouns, but always gone by 80ms). Further evidence of root priming's comparable robustness comes from its imperviousness to semantic opacity and allomorphic variation.

In an earlier cross-modal priming study, Boudelaa and Marslen-Wilson (2000) examined the conditions under which root and pattern priming would obtain. They compared root priming between pairs of words where the morphological relationship was semantically transparent (dirasa-madrasa, lesson-school) and pairs of words where the relationship was opaque (muqtaniʕ- muqannaʕ, satisfied-masked). Root priming was found to obtain despite opaque semantics. This finding was in direct contrast with their results for pattern priming, which was only observed when both the prime and target patterns carried congruent syntactic information.

Regarding allomorphic variation, Boudelaa and Marslen-Wilson's (2004) masked priming study compared priming between "strong" roots with transparent phonology, with priming between allomorphic variations of "weak" roots (roots with one letter that changes depending on the phonetic environment the pattern puts it into). For example, w-f-q is a weak root. Its first letter is transparent in the surface form waafaqa, 'agreed' but it appears as a [t] in the surface form ittifaaq, 'agreement'. Root priming was found to obtain despite such allomorphy, (evidence, the authors argued, of its phonologically abstract nature). Pattern priming, on the other hand, was only observed when both the prime and target patterns had intact CV skeletons, undisrupted by allomorphy.

The authors explained the differences between root and pattern priming across these studies by appealing to differences in their functional and distributional properties. Patterns are too productive and the syntactic roles they signify too general to efficiently pare down competitors during lexical access. Because roots are less productive than patterns and are more focused in terms of their semantic features, it makes sense for an Arabic speaker's lexicon to be organized around them. (It is for this same reason that Arabic dictionaries are organized by roots.) In order to test their hypothesis that the features of roots drive lexical access in Arabic, Boudelaa and Marslen-Wilson (2011) designed a study to determine the effects of the productivity of both roots and patterns on pattern priming. In a masked priming experiment, they varied the productivity of the roots and the patterns in prime-target pairs whose patterns overlapped and found that root productivity alone determined the strength of the pattern priming.

In conclusion, primed LDT research in Arabic suggests that root and pattern morphemes are independent at some level of mental representation and that identifying these

morphemes is an obligatory part of lexical access. Of the two kinds of morphemes, root priming appears to be the faster, more robust process, whereas pattern priming is slower and more easily interfered with.

2.5.2 L2 Processing of Arabic Derivational Morphology at the Lexical Level

Evidence of morphological priming in L2 learners of Arabic comes from Freynik, Gor and O'Rourke (submitted), who adapted Boudelaa and Marslen-Wilson's (2000) methodology to test whether L2 learners of Arabic whose L1 is English would show speeded RTs to target words that were preceded by primes that shared the same root morphemes, particularly when the prime-target relationship was semantically opaque. The L2 learners showed significant root priming, both in semantically transparent and semantically opaque conditions, suggesting that L2 learners are able to decompose Arabic words into their constituent morphemes and make use of them during lexical access, in spite of their discontinuous structure and inconsistent semantic contribution. L2 learners' sensitivity to Arabic derivational morphology is interesting because, while the Combinatorial Entries Hypothesis could arguably account for L2 priming observed between Germanic derived words and their stems, it is hard to see how it should account for the effects observed in Arabic.

Between the greater productivity of Arabic derivational morphology, the obligatoriness of decomposition and the fact that the focus was priming between two derived forms (as opposed to between a derived form and a constituent stem), the requirements of processing Arabic derivational morphology are comparable to those of processing inflectional morphology in Indo-European, at least at the lexical level. As Marslen-Wilson (2007) explains,

> West-Germanic languages like English, Dutch or German may exemplify the kind of situation sketched by Clahsen et al. where all derived forms have, by definition, a "lexical entry" in the neurocognitive language system, but where only a subset of these, based on transparent and productive word-formation processes, are stored in a decomposed and combinatorial format. It would be only this subset then, that could support the kinds of lexicon-wide representational linkages that are detected up in overt priming tasks - and perhaps also the same subset that accounts for most of the variance in studies of morphological family size. For a language like Arabic, in contrast, it is possible that all complex forms are stored in a morphologically decomposed format, so that there are not the same variations in accessibility to a word's morphemic components as a function of priming task and semantic transparency. But a great deal of further research is needed to flesh out these speculations (Marslen-Wilson, 2007, p.188 - 189).

Indeed, this is an empirical question; if derivational and inflectional Arabic morphology are handled similarly at the lexical level, similar priming should be observable between derived and inflected forms during lexical access. Differences should emerge at the sentence processing level, where the two kinds of morphology are functionally distinct.

2.5.3 L1 Processing of Semitic Inflectional Morphology at the Sentence Level

While no investigations of inflectional morphology at the sentential level have been carried out in Arabic, one has been conducted in Hebrew, a Semitic language whose system of morphology is similar. Deutsch (1998) used eye-tracking software to compare participants' reading times for grammatical sentences to their reading times when sentences included errors in gender or number agreement between subject and verb across two conditions: a short-distance condition in which the subject and verb were adjacent, as in example (2.6a), and a long-distance condition in which the subject and verb were separated by five intervening words as in example (2.6b).

2.6a *Hashoter divach ki mekhoniyot (fem., pl. - cars) nigneva (fem., sing. - had been stolen)

beshaa chamesh lifnot boker.

*The policeman reported that cars had been stolen at five o'clock in the morning.

2.6b *Hashoter diavach ki mekhoniyot (fem., pl. - a car) mishtara mehadegem hayafe vehachadish beyoter nigneva (fem., sing. - had been stolen) beshaa chamesh lifnot boker.

*The policeman reported that police cars of the nicest and most recent model had been stolen at five o'clock in the morning.

Deutsch found that participants exhibited longer reading times for non-agreeing verbs only in the short-distance condition. Deutsch interpreted this result as evidence that syntactic features are accessed quickly but fleetingly, and that they are soon replaced by the semantic features, which are accessed subsequently. Thus, what research exists on Semitic inflectional morphological processing in sentential contexts suggests that the same patterns observed for other languages also hold for Semitic languages. That is to say, native speakers tend to show slower reading times following agreement errors in inflectional morphology.

3 The Current Study

To recap, lexical level investigations into L2 morphological processing have suggested that L2 learners may store and access derivational morphology in a more native-like way than they do inflectional morphology (Silva and Clahsen, 2008; Diependaele et al., 2011; Kim et al., 2011). One explanation that has been offered for this pattern cites differences between the lexical representations of derived forms and inflected forms as the source of their relative acquirabilities. Specifically, the Combinatorial Entries Hypothesis holds that derivational morphology does not require decomposition in the same way that inflectional morphology does. While this analysis of Indo-European derivational morphological representations may be accurate, Arabic derivational morphology appears to require full decomposition during lexical access. Nevertheless, L2 learners of Arabic appear to be able to process Arabic derivational morphology in a way similar to native speakers, at the lexical level. This casts doubt on the source of the discrepancies between derivational and inflectional morphemes in second language acquisition. The goal of the current study is to compare the two kinds of morphology directly, and to clarify what makes derived forms easier to acquire.

The current study is an examination of how inflectional and derivational morphology are processed by L2 learners at both the lexical and the sentential level (and how L2 behavior compares to that of native speakers across these conditions). As the previous sections have explained, Arabic verbal morphology allows for inflectional and derivational manipulations to the same roots, which can be compared to one another at the lexical

level in terms of their decomposability. Further, both derivational and inflectional manipulations to Arabic verbs can result in sentence-level anomalies, such that it is possible to compare them at the sentence level, without using pseudowords.

3.1 Research Questions

The research questions that the current study aims to address are the following:

1. Are L2 learners more sensitive to Arabic derivational morphology than to Arabic inflectional morphology at the lexical level?
2. Is their morphological sensitivity limited to the lexical level or can L2 learners make use of this morphological information during sentence processing (i.e., how "deep" and automatic is their knowledge of this morphology)?
3. How does automatic or integrated L2 knowledge of morphology compare to explicit, conscious L2 knowledge of morphology?

4 Methods

The section that follows describes the current study's use of a lexical decision task, an acceptability-judgment task, and a self-paced reading task to triangulate a picture of L1 and L2 Arabic learners' processing of derivational and inflectional morphology at the lexical and sentential levels.

4.1 Participants

Forty-four L2 learners and 34 native speaker participants were recruited from Arabic language programs and Arab student associations at the University of Maryland and other American universities (including Georgetown University, American University, the University of Texas at Austin, Penn State University and Brigham Young University), by posting flyers on campuses and emailing student listservs. A short questionnaire was given to volunteering participants to determine that they were either (a) native-speakers who were born in and had lived in Arabic-speaking countries for at least the first 10 years of their lives, or (b) L2 learners of Arabic who had studied Arabic for at least 2 years, and had not been exposed to the language before high school.

Among the L1 participants, the average age was 30.9 years, with a minimum of 21 and a maximum of 42. There were 12 participants from Egypt, 7 from Jordan, 4 from Lebanon, 3 from Iraq, 3 from Morocco, and 1 participant from each of the following countries: Palestine, Saudi Arabia, Tunisia, Yemen, and Libya. Among the L2 learner participants, the average age was 26.4 years, with a minimum of 19 and a maximum of 37. L2 learner participants had an average of 4.1 years of formal Arabic study, with a minimum of 2.5

years and a maximum of 7 years. They had spent an average of 1.75 years living in an Arabic-speaking country, with a minimum of 0 years and a maximum of 6 years. Zero to 6 years of immersion is a wide spread, but both the minimum and maximum points on this continuum were outliers. The majority (24) of the 44 total L2 participants had between 0.5 and 1 year experience living in an Arabic speaking country. Two participants had never lived in an Arabic speaking country, six had spent 2 years in one, and eight had spent 3 years. One participant had spent 4 years, two had spent 5, and one had spent 6.

4.2 Experiment 1 – Primed Lexical Decision

4.2.1 Task

Experiment 1 used cross-modal priming to investigate the role inflectional and derivational morphemes play in L2 Arabic learners' lexical processing. The relation between prime and target was manipulated across six conditions.

4.2.2 Conditions

The first condition, derivational, tests for morphological priming between verbs that are derived from the same triliteral root. This is the Arabic analog of the morphological priming found in Indo-European languages between derived words that share a stem.

The second condition, inflectional, tests for morphological priming between an inflected verbal prime and a target that consists of the same verb's unmarked (base) form. This is the Arabic analog of the morphological priming found between inflected and stem forms such as walk-ed and walk in English.

The last three conditions are controls. The third condition, phonological, provides a baseline for phonological priming (all primes in this condition share at least 3 phonemes

with the target, but they are not all root letters). Phonological overlap between the derivational and inflectional conditions is also controlled in terms of surface similarity, such that the phonological condition is an equally adequate control for both morphologically related conditions. Specifically, if the derivational prime for a given target has the same onset as that target, then the inflectional prime likewise has the same onset, as does the phonological control. Conversely, if one of these conditions has a different onset than the target, all three have a different onset. All three conditions likewise share the same number of phonemes overall with the target on average, give or take no more than one phoneme in particular.

The fourth condition, semantic, represents a baseline for semantic priming in the absence of a morphological relationship (all primes in this condition were judged by 5 native speakers to have an average semantic association of 7 or higher on a 9 point scale).

The fifth condition, unrelated, provides a baseline for participants' RTs when the prime bears no relationship to the target at all (that is, it shares no more than one consonant, and were judged by native speakers to have a semantic association of 3 or lower on a 9 point scale).

4.2.3 Design

The experimental items come from a master list constructed of 80 target words. Each target word has a corresponding set of five potential prime words: one for each of the five conditions. Using a Latin Square design, five experimental lists are created from this master list, such that every target word appears with a prime word from a different priming condition in each of the lists. Table 4.1 demonstrates this paradigm.

Table 4.1 Example sextets

Deriv	Infl	Phon	Sem	Unrel	Target
khawwafa scare	*khaafuu* fear(pl)	*khalafa* succeed	*ruu3a* frighten	*qaala* say	*khaafa* fear
aTa33ama feed	*yaT3amu* eats	*taTaba3a* prints	*akal* food	*Daraba* hit	*Ta3ama* Eat
‘*asqaTa* drop	*yasqaTu* falls	*taqaTTa3a* cut	*waqa3a* fall	*mathalan* approximate	*saqaTa* fall
ta3arrafa meet	*ya3rafu* knows	*ta3afaa* forgive	*darasa* learn	*nazala* descend	*3arafa* know
‘*afhama* explain	*yafhamu* understands	*muhimma* important	*waDaHa* clarify	*DaHaka* laugh	*fahama* understand

Each list contains 16 items in each condition. Experiment 1 is loosely modeled after Boudelaa and Marslen-Wilson (2000), who found significant effects with only 6 and 8 items per condition; however, L2 learners are not necessarily expected to know all of the target words, such that the best strategy is to cast a relatively wide net in order to gather a useful sample of valid trials (see the description of the post-LDT Vocabulary Survey below for further discussion of how valid trials are determined). Moreover, L2 behavioral data tends to exhibit a broader range of variation than L1 data does (Eckman, 1994; Tarone, 1988).

A one-to-one word-to-nonword ratio balances the lexical decision task and prevents participants from developing a guessing strategy; thus, 80 nonwords were created. 40 were created by combining nonexistent triliteral roots (e.g., b-k-t) with existing word patterns, while 40 were created by combining existing roots with nonexistant patterns. This method of nonword creation was chosen because rejecting nonwords with nonexistent roots or patterns should be a straightforward and comparably easy task (compared to, say, rejecting nonwords that are composed of licit roots and licit patterns whose combination is not found in the lexicon. Rejecting such possible-but-nonexistent pseudowords has been shown to require not just morphological decomposition but also the further stages of re-composition and checking (Taft, 2004).

In order to prevent learners from associating phonological similarity between prime and target with the target's lexical status, 48 of the nonword trials were preceded by prime words that shared at least two of their consonants (such that, for both the nonword and the real word trials, 40% of the trials exhibited no phonological relationship between the prime and the target. That is to say, phonological relatedness did not statistically predict lexical status across items).

4.2.4 Vocabulary Post-test

After completing all 3 experimental tasks, L2 participants completed a Vocabulary Survey during which they were given a list of all the real Arabic words that appeared in the lexical decision task, and were asked to write an English translation for each. Their performance on this Vocabulary Survey was used to filter the lexical decision items for analysis; if a participant could not translate both words in a prime-target pair, that item

was excluded from analysis. This vocabulary measure was furthermore used, in addition to self-reporting, as an estimate of L2 learners' Arabic proficiencies.

4.2.5 Analysis

Once the items are filtered for accuracy and vocabulary knowledge, response times to the remaining trials will be inspected visually in bar graphs and boxplots to look for trends. They will then be subjected to analyses of variance. To compare the effects of the conditions on reaction times, four 2X2 ANOVAs are planned comparing mean RTs in each of the priming conditions (Derivational, Inflectional, Phonological and Semantic) to the Baseline condition, with Language Group (L1 vs. L2) as the between-subjects factor in all three. Two additional 2x2 ANOVAs will compare mean RTs in each of the morphological priming conditions (Derivational and Inflectional) to the Phonological condition to establish whether morphological priming is distinct from phonological overlap. Simple comparisons will be used to investigate any significant interaction effects.

4.2.6 Predictions

Based on the findings in Freynik, Gor and O'Rourke (submitted), L1 and L2 participants alike are expected to show speeded RTs in the Derivational priming condition, relative to the Phonological and Semantic control conditions. Based on studies of inflectional morphology in other languages (e.g., Silva and Clahsen, 2008; Gor and Cook, 2010), L1 participants are further expected to show speeded RTs in the Inflectional priming condition.

If the Combinatorial Entries Hypothesis is correct that L2 learners store derived forms in a more native-like way than they do inflected forms, L2 learners should show significantly faster RTs in the Derivational priming condition than in the Inflectional priming condition.

If, conversely, the L2 learners show comparable degrees of priming in the Derivational and Inflectional conditions, this would constitute evidence that L2 learners are comparably capable of storing and decomposing both kinds of morphologically complex forms into their constituent sublexical structures. Even if L2 learners are capable of decomposing inflected verbs during lexical access, however, this does not clarify whether inflectional or derivational morphology is comparably easier to make use of in sentential contexts, either due to the features they encode (Uninterpretable Features Hypothesis) or due to the sentence level dependencies they signal (Sentence Level Dependencies Hypothesis). Shedding light on these questions is the purpose of Experiment 2.

Table 4.2 Summary of Experiment 1 Predictions

	Derivational	Inflectional
Combinatorial Entries Hypothesis	priming	no priming
Uninterpretable Features-Hypothesis	no claims	
Sentence-Level Dependencies Hypothesis	no claims	

4.3 Experiment 2 - Self-Paced Reading

4.3.1 Task

If Experiment 1 was designed to probe L2 learners' online sensitivity to Arabic inflectional and derivational morphology at the lexical level, experiment 2 is designed to measure L2 participants' automatized command of inflectional and derivational morphology during sentence processing. To this end a self-paced reading task was adapted to measure learners' sensitivity to the different kinds of morphological errors during sentence reading. Critical items consist of sentences which always become ungrammatical at the matrix verb. Recall from section 2.4 above that the structure of Arabic verbs allows for the direct comparison of inflectional and derivational violations in the context of the same verbal root forms.

4.3.2 Conditions

The first condition, Baseline, is an acceptable Arabic sentence.

4.1a Fasara al-kitaab an al-temthaal **ta'asasa** fii hatha al-makaan mundhu adat senawaat.

explained the.book that the.monument **was.built** in this the.place ago number years

The book explained that the monument **was built** in this place several years ago.

In the second condition, Derivational, the anomaly results from the wrong derivational verbal template being applied to an otherwise appropriate verbal root.

4.1b Fasara al-kitaab an al-temthaal **'asasa** fii hatha al-makaan mundhu adat senawaat.

explained the.book that the.monument **built** in this the.place ago number years

The book explained that the monument **built** in this place several years ago.

In the third condition, Inflectional, the violation always arises from the wrong number agreement marking on the verb.

4.1c. Fasara al-kitaab an al-temthaal **ta'asasuu** fii hatha al-makaan mundhu adat senawaat.

explained the.book that the.monument **were.built** in this the.place ago number years

The book explained that the monument **were built** in this place several years ago.

Because the anomalies that arise in the derivational condition have to do with the subjects being inappropriate agents for the verb forms in question, a semantic control condition is included to gauge participants' willingness to judge sentences unacceptable when they are anomalous on semantic grounds alone (i.e., when that sentence cannot be salvaged by substituting different derivational morphology). In the semantic control condition, the anomaly arises because the verb names an action that the subject could not reasonably perform.

4.1d. Fasara al-kitaab an al-temthaal **'asafa** fii hatha al-makaan mundhu adat senawaat.

explained the.book that the.monument **regretted** in this the.place ago number years

The book explained that the monument **regretted** in this place several years ago.

4.3.3 Design

The critical materials include 60 sets of sentences across 4 conditions like the ones described above (3 error conditions and a grammatical condition). Using a Latin square rotation, 4 lists were created with 15 items in each condition such that no single participant will read 2 versions of the same sentence.

In addition, 116 grammatical distractor sentences were added to tip the proportion of grammatical to ungrammatical sentences to nearly 3:1 (in total there were 45 erroneous sentences and 131 correct sentences in each list), as well as to vary the structure of the target sentences so that participants would be less likely to develop strategies. In order to ensure that participants read the sentences for meaning, half of all the sentences were followed by comprehension questions with yes/no answers.

4.3.4 Analysis

To compare the effects of the conditions on reading times, at each of four regions of interest (the precritical region just before the error word, the critical region where the error occurs, spillover region 1 immediately following the critical region, and spillover region 2, two words after the critical region), three 2X2 ANOVAs are planned comparing mean RTs in each of the critical error conditions (Derivational, Inflectional and Semantic) to the Baseline condition, with Language Group (L1 vs. L2) as the between-subjects factor in all three. Simple comparisons will be used to investigate any significant interaction effects.

4.3.5 Predictions

L1 learners are expected to have slower reading times for the words that follow the bolded word in all the error conditions described. Precedent for L1 learners exhibiting slower reading times following SV agreement errors comes from Pearlmutter et al. (1999) and the L1 controls in Jiang (2004). Precedent for L1 learners reading more slowly following inappropriate derivational morphology comes from Clahsen and Ikemoto (2012).

Regardless of how automatized (or not) their morphological processing is, L2 learners are expected to exhibit slowdowns in the region following the semantic anomaly in the semantic condition. If L2 learners have native-like sensitivity to both kinds of morphological anomalies, then they would be expected to exhibit similar slow-downs following errors in the inflectional and derivational conditions as well. However, if familiarity-from-the-L1 is the relevant factor for determining which morphological features L2 learners come to interpret in a native-like manner, then L2 learners whose L1 is English should exhibit slowdowns in the region following the verb in the inflectional error condition, with the number agreement errors, because English marks number agreement between subjects and verbs. Conversely, the verbal derivation paradigm of Arabic is unlike English verbal derivational morphology. This is not to say that the event structures themselves are unfamiliar to English speakers. It is likely that the event structures that correspond to the ten verb forms find some expression in all languages. Doron and Hovav (2009), for instance, have argued that the reflexivity which is realized in Semitic languages with a verbal pattern is realized in Romance languages with a clitic, and in English with a null morpheme. So, it is not the case that the semantics of

reflexivity (or causativity, or anti-passivity) are unfamiliar to English speakers, any more than the semantics of plurality are unfamiliar to Chinese speakers. What is unfamiliar to English speakers is the correspondence between the Arabic verbal patterns and the event structures they signal (recall the discussion of root and pattern interleaving in Section 2.4 above). In this sense, the Uninterpretable Features Hypothesis predicts that L2 learners should not be sensitive to morphological errors of this kind. That is, they should not exhibit slowdowns in the regions following the anomaly in the derivational error condition.

Conversely, if the difficulty of interpreting a given morpheme depends on the nature of the dependency it signals, then the derivational error condition should be the easier one for L2 learners to interpret, because, as explained in section 2.4.3 above, the verbal derivational morphology in the derivational error condition determines the verb's thematic roles, and this information is stored in the lexicon. The appropriateness of a thematic role can be gauged without articulating syntactic structure (e.g., "the monument builds…" is bad in the same way that "the monument regrets" is bad, and in a different way than "the monument build…" is bad). Thus, the Sentence Level Dependencies Hypothesis predicts that L2 learners should exhibit slowdowns in the regions following the anomaly in the derivational error condition and not in the region following the anomaly in the inflectional error condition. These predictions are summarized in Table 4.3 below.

Table 4.3 Summary of Predictions for Experiment 2

	Derivational	Inflectional	Semantic
Combinatorial Entries Hyp	no specific claims about sentence processing		
Uninterpretable Features Hyp	no slow-downs	slow-downs	slow-downs
Sentence-Level Dependencies	slow-downs	no slow-downs	slow-downs

4.4 Experiment 3 – Acceptability Judgment Task

4.4.1 Task

Experiment 3 is intended to measure L2 learners' offline knowledge of Arabic derivational and inflectional morphology at the sentence level. The goal of this experiment is to establish which kinds of morphological errors L2 learners may be aware of under the most favorable circumstances, that is, when they are under no time pressure and when they are instructed to attend consciously to form as well as meaning. Such a task cannot distinguish whether this knowledge is automatized or merely consciously controlled.

To this end, an acceptability judgment task (AJT) was designed wherein the critical sentences are either correct or anomalous according to four critical conditions and seven filler conditions. The four critical conditions correspond to the critical conditions in the self-paced reading task: baseline, derivational, inflectional and semantic. Because the AJT was devised for comparison with the SPR in order to shed light on how different

task demands might affect learners' sensitivities to the same kinds of errors, the same subject-verb pairs occur in the critical AJT items as did in the critical SPR items.

The filler items were divided into three additional inflectional conditions (feminine gender agreement, first person agreement, and second person agreement), as well as four derivational conditions (causative, anticausative, passive, and reflexive). These filler conditions were designed to probe learners' knowledge of additional categories that were not possible to include in the self-paced reading task.

4.4.2 Conditions

As mentioned above, the four critical conditions correspond to the ones in the self-paced reading task. The first condition, Baseline, is an acceptable Arabic sentence.

4.2a Fii gharb al-bilaad, al-baarid **Hadhara** al-naas min al-shitaa' al-qaadim
in west the.country the.cold **warned** the.people from the.winter the.coming
In the west of the country, the cold **warned** the people of the coming winter.

In the second condition, Derivational, the anomaly results from the wrong derivational verbal template being applied to an otherwise appropriate verbal root.

4.2b Fii gharb al-bilaad, al-baarid **Haadhara** al-naas min al-shitaa' al-qaadim
in west the.country the.cold **was.careful** the.people from the.winter the.coming
In the west of the country, the cold **was careful** the people of the coming winter.

In the third condition, Inflectional, the violation arises from the wrong number agreement marking on the verb.

4.2c Fii gharb al-bilaad, al-baarid **Hadharuu** al-naas min al-shitaa' al-qaadim

in west the.country the.cold **warned.pl** the.people from the.winter the.coming

In the west of the country, the cold **warned(pl)** the people of the coming winter.

Because the anomalies that arise in the derivational condition have to do with the subjects being inappropriate agents for the verb forms in question, a semantic control condition is included to gauge participants' willingness to judge sentences unacceptable when they are anomalous on semantic grounds alone (i.e., when that sentence cannot be salvaged by substituting different derivational morphology). In the semantic control condition, the anomaly arises because the verb names an action that the subject could not reasonably perform.

4.2d Fii gharb al-bilaad, al-baarid **Hasada** al-naas min al-shitaa' al-qaadim

in west the.country the.cold **envied** the.people from the.winter the.coming

In the west of the country, the cold **envied** the people of the coming winter.

In the first filler condition, Gender Agreement, the violation arises from the wrong gender agreement marking on the verb.

4.3a Al-ghaTla **kalafat** al-sherika qurd kebiir min al-amwaal.

the.mistake.fem **cost.fem** the.company amount large from the.money

The mistake(fem) **cost(fem)** the company a large amount of money.

4.3b Al-ghaTla **kalafa** al-sherika qurd kebiir min al-amwaal.

the.mistake.fem **cost.masc** the.company amount large from the.money

The mistake(fem) **cost(masc)** the company a large amount of money.

In the second filler condition, First Person Agreement, the violation arises from the wrong person agreement on the verb.

4.4a Akhii saqaTa haatifii walakin-ii **3adhartuhu** ba3da dhalik.
brother.my dropped phone.my but.me **forgave.1st.him** after that
My brother dropped my phone but **I forgave him** later.

4.4b Akhii saqaTa haatifii walakin-ii **3adharahu** ba3da dhalik.
brother.my dropped phone.my but.me **forgave.3rd.him** after that
My brother dropped my phone but **I forgave him** later.

In the third filler condition, Second Person Agreement, the violation likewise arises from incorrect person agreement on the verb.

4.5a ASdiqaaik istaqaaluu andama badaa' al-thelj, walakinak **Samadta**.
friends.your quit.pl when began the.snow but.you **perservered.2nd**
Your friends quit when it began to snow, but you **perservered.**

4.5b ASdiqaaik istaqaaluu andama badaa' al-thelj, walakinak **Samada.**
friends.your quit.pl when began the.snow but.you **perservered.3rd**
Your friends quit when it began to snow, but you **perservered.**

In each of the four derivational filler conditions, which are discussed next, the anomalies result from a wrong derivational verbal template being applied to an otherwise

appropriate verbal root. The first condition, causative form, involves using a causative form verb with a subject that is semantically implausible as its agent.

4.6a ba3ada min al-naas fii ruusiya Sawarat al-niizik 3andama **Saqata**.
some from the.people in Russia filmed the.meteor when it **fell**
Some of the people in Russia filmed the meteor as it **fell**.

4.6b ba3ada min al-naas fii ruusiya Sawarat al-niizik 3andama **aSaqata.**
some from the.people in Russia filmed the.meteor when it **fell.transitive**
Some of the people in Russia filmed the meteor as it **fell(transitive).**

In the second derivational condition, anticausative form, the anomaly arises from a similar semantically inappropriate subject-verb combination wherein the verb's anticausative form makes it something the subject cannot do.

4.7a Ma kaana ijaaba, hakadha al-zaa'ir **fataHa** al-baab.
Not was answer, thus the-visitor **opened** the-door
There was no answer, so the visitor **opened** the door.

4.8b Ma kaana ijaaba, hakadha al-zaa'ir **infataHa** al-baab.
Not was answer, thus the-visitor **opened.by.itself** the-door
There was no answer, so the visitor **opened(by itself)** the door.

In the third derivational condition, reciprocal form, the anomaly arises because the verb is in the reciprocal form, but the subject makes an inappropriate agent for a reflexive action.

4.9a Andama min al-laazim, al-muzaare3 **yedrabu** al-Himaar.

When from the-necessary, the-farmer **hits** the.donkey

When necessary, the farmer **hits** the donkey.

4.10b Andama min al-laazim, al-muzaare3 **yetadaarabu** al-Himaar.

When from the-necessary, the-farmer **hits.each.other** the.donkey.

When necessary, the farmer **hits(reciprocal)** the donkey.

In the fourth derivational condition, the anomaly arises because the verb is in the passive form, but the subject makes an inappropriate argument for the passive form of that action.

4.11a Shakhsan maa **adkhala** risaala bi-fetHa al-bariid.

Person what **inserted** letter with.opening the.mail

Someone **inserted** a letter into the mail slot.

4.11b Shakhsan maa **tadkhala** risaala bi-fetHa al-bariid.

Person what **was.inserted** letter with.opening the.mail

Someone **was inserted** a letter into the mail slot.

4.4.3 Design

The critical materials include 60 sets of sentences across 4 conditions like the ones described above (3 error conditions and a grammatical condition). Using a Latin square rotation, 4 lists were created with 15 items in each condition such that no single participant will read 2 versions of the same sentence.

The filler items come from a master list of 110 pairs of sentences across the seven filler conditions, where each pair consists of one acceptable and one unacceptable version of the same sentence. Roughly 16 pairs were created for each condition (some filler conditions had 14 and some had 18). From this master list, two complementary experimental lists were made, with 8 acceptable and 8 unacceptable sentences per condition, such that each experimental list consisted of 110 sentences total, with no participant seeing two versions of the same sentence. Additionally there were 30 correct filler sentences added to balance out the proportion of acceptable to unacceptable sentences to 50:50.

Sentences are presented via Ibex Farm using an AJT script that presents each item in a single line of text. Participants respond to each sentence using the Right or Left Control Key to indicate whether the sentence is acceptable (Right for 'Yes' and Left for 'No'). During a 10 sentence practice session (with feedback provided), participants are familiarized with this task (no feedback is provided during the rest of this task, i.e., the non-practice phase).

4.4.4 Analysis

To compare the effects of the conditions on accuracy scores, three 2X2 ANOVAs are planned comparing mean accuracy scores in each of the critical error conditions (Derivational, Inflectional and Semantic) to the Baseline condition, with Language Group (L1 vs. L2) as the between-subjects factor in all three. Simple comparisons will be used to investigate any significant interaction effects.

4.4.5 Predictions

L1 participants are expected to recognize anomalous sentences across all conditions with a high (i.e., ceiling) degree of accuracy. Substantial precedent for L1 learners' accuracy in judging faulty SV agreement ungrammatical comes from the L1 controls in studies like Johnson & Newport (1989) and Jiang (2004). It is likewise probable that L1 learners will judge sentences with inappropriately derived verbs to be unacceptable, just as L1 participants in Clahsen & Ikemoto's (2012) study rated sentences with inappropriately derived nominal to be unacceptable.

Because an AJT is amenable to explicit knowledge of grammaticality (and in this case, semantic plausibility), and all of the anomalous sentence conditions involve linguistic phenomena that are addressed in the L2 classroom, it is expected that performance on this task will correlate with L2 learners' Arabic proficiency in all conditions. As proficient L2 Arabic learners were sought to participate in the current study, it is expected that their accuracy on the AJT will reflect their high proficiency.

The Combinatorial Entries Hypothesis is a claim about the way derivational as opposed to inflectional morphology is stored and accessed in the L2 mental lexicon; it is not a claim about how that morphology is interpreted offline in sentence contexts. Thus, this hypothesis makes no specific claims about the AJT task.

The Uninterpretable Features Hypothesis, by contrast, is a claim about L2 learners' sensitivity to different morphemes during sentence processing, depending on the features those morphemes signal. It suggests that L2 learners can come to process morphological information in a more native-like and automatic way when the morphemes in question

encode features that are familiar from the L1. This hypothesis, then, does not imply that L2 learners should not be able to gain explicit knowledge about any kind of morphology at all, and thereby make accurate judgments about it on the AJT. That said, if morphology with L1-familiar features can come to be processed in a more native-like, automatic way, while other morphology is consciously and effortfully processed, then L2 learners should more reliably make accurate judgments about the morphology they process automatically than about the morphology they have to consciously consider. In this experiment, subject-verb number agreement, as a feature, is familiar from English, while subject-verb gender agreement is not. For this reason, if there is a discrepancy in accuracy between the conditions (as opposed to a ceiling effect), the Uninterpretable Features Hypothesis would still predict that judgments about number agreement should be more accurate than judgments about gender agreement. Judgments about number agreement should likewise be more accurate than judgments about derivational morphology, as the correspondence between the Arabic verbal patterns and the event structures they signal in those conditions with derivational violations are not familiar from English morphology.

The Sentence-Level Dependencies Hypothesis is a claim about L2 learners' ability to represent and interpret sentence-level dependencies during sentence processing. According to this hypothesis, the relative difficulty of interpreting a given sentence-level dependency corresponds to the distance that dependency spans (e.g., how many words intervene between agreeing constituents) and its structural complexity (e.g., does it span an embedded clause?). Clahsen et al. (2010) explained the way different kinds of dependencies might tax L2 learners' processing to different degrees in their review of Sato's (2007) results, wherein Sato found that L2 learners from three different L1

backgrounds (German, Japanese, and Chinese) were all comparably more accurate when it came to making judgments about English pronominal case than when it came to making judgments about English subject-verb agreement. Clahsen et al. (2010) notes,

> Another difference between the two phenomena under investigation is that SV agreement dependencies span the entire clause (and thus require comparatively complex structural scaffolding), whereas the objective case is assigned locally within the verb phrase. Sato's results may thus reflect learners' relatively greater difficulty establishing clause-level morphosyntactic dependencies under processing pressure. (Clahsen et al., 2010, p. 37)

By this logic, interpreting subject-verb agreement involves accurately establishing a clause-level structural dependency, whereas assessing the appropriateness of a thematic relation between, for example, a given subject and a causative predicate, can be accomplished by virtue of the argument structure assigned by the verb's lexical entry. Assessing an anomaly of this sort in the derivational conditions should involve the same kind of structural difficulty implicit in assessing the anomaly in the semantic error condition, (e.g., in saying, "the fire asks…"). In this sense, this kind of dependency should be the easier one for second language learners to interpret, so that if there is a discrepancy in accuracy between conditions, the Sentence Level Dependencies Hypothesis would predict that judgments about derivational morphology (which in the current experiment determines thematic roles) should be more accurate than judgments about inflectional morphology (which in the current study corresponds to subject-verb agreement). As in the discussion of the Uninterpretable Features Hypothesis' predictions, however, it must also be noted here that an AJT is always amenable to conscious monitoring with explicit knowledge, and that L2 learners with accurate explicit knowledge about Arabic morphology should be able to give accurate judgments about

even those forms that are difficult to process, and which they may have not automatized, as well as to map out those dependencies that they might fail to interpret under time pressure. Table 4.4 summarizes the L2 predictions for each condition of Experiment 3, then, with the caveat that ceiling accuracy is possible for all conditions about which L2 participants have explicit knowledge.

Table 4.4 Summary of Predictions for Experiment 3

<table>
<tr><td rowspan="2"></td><td colspan="4">Derivational</td><td colspan="2">Inflectional</td><td rowspan="2">Semantic</td></tr>
<tr><td>Caus</td><td>Anticaus</td><td>Recip</td><td>Passive</td><td>Number</td><td>Gender</td></tr>
<tr><td>Combinatorial Entries Hyp</td><td colspan="7">no specific claims about offline sentence processing</td></tr>
<tr><td>Uninterpretable Features Hyp</td><td colspan="4">lower accuracy</td><td>higher accuracy</td><td>lower accuracy</td><td>Ceiling</td></tr>
<tr><td>Sentence-Level Dependencies</td><td colspan="4">higher accuracy</td><td colspan="2">lower accuracy</td><td>Ceiling</td></tr>
</table>

4.5 Procedure

Participants completed the experimental tasks by way of Ibex's ("Internet Based EXperiments") remote testing capability[1]. Before participating in the experiment, participants were instructed to read the consent form (requirement to sign the consent form was waived by the Institutional Review Board in light of remote testing). They were then instructed to complete the language history questionnaire (attached as Appendix A), which asks for such information as languages spoken, the ages at which they began to learn each, years of formal instruction, approximate percentages of time spent using each

[1] For more information on Ibex's remote testing capability, see http://spellout.net/latest_ibex_manual.pdf

language during different periods of their lives, self-reported proficiencies in different modalities (e.g., listening, writing), gender, age, nationality and handedness.

4.5.1 Lexical Decision Task

Participants are then instructed to click a link which will start the lexical decision experiment; at the beginning of the lexical decision experiment, the written instructions for the task appear on the screen. After reading the instructions, participants begin a short practice phase (15 prime-target pairs) with feedback (the actual task does not involve feedback), after which they are given the opportunity to begin the task or restart at the beginning of the instructions. A break is provided at the midpoint during the experiment.

The primes for Experiment 1 are spoken by a male native speaker of Arabic and digitally recorded in wav file format. At the beginning of each trial, a fixation mark appears in the center of the computer monitor for 200ms and stays onscreen while the wav file of the auditory prime word is heard over the headphones. At the offset of the prime word, the written target word is presented in the same location as the preceding fixation point in a 30-point traditional Arabic font size. The target stays on the computer monitor for 1000ms. The starting point for measuring reaction times begins with the onset of the target word. After the first 1000ms, the target word will disappear, leaving a blank screen. The trial was supposed to time out after 3000ms, if no response was given. A coding error in the items file, however, resulted in an error that caused each trial to last until the participant responded. Section 6 discusses this issue in greater detail below.

Participants respond by pressing the Right or Left Control Key on the keyboard (Right was for real word and Left was for non-word; participants are given 15 practice trials to

get accustomed to the timing and response keys). The presentation of the stimuli and the measuring of the reaction times are handled by the Ibex Farm system.

After the cross-modal priming task, L2 participants are instructed to take a break before moving on to the Self-Paced Reading Task.

4.5.2 Self-Paced Reading Task

Following the Lexical Decision Task, participants are instructed to click a second link, which starts the self-paced reading task. At the beginning of the self-paced reading task, the written instructions for the task appear on the screen. After reading the instructions, participants begin a short practice phase (10 example sentences, 5 of which are followed by comprehension questions and feedback – the task involves no feedback once practice is over), after which they are given the opportunity to begin the task or restart the instructions. A break is provided at the midpoint during the task.

The sentences for the self-paced reading task are presented in a 30 point traditional Arabic font, in a single line of text, which is initially masked by horizontal dashes. At the beginning of each trial, the whole sentence appears as a series of horizontal, word-length dashes. Participants press the space bar to unmask the sentence with a "moving window" that displays one word at a time. After the sentence-final word, pressing the space bar again leads to a 500ms blank screen. If there is a comprehension question for the sentence, it will appear next. Participants respond to the question by pressing the Right or Left Control Key on the keyboard (Right for 'Yes' and Left for 'No'). The presentation of the stimuli and the measuring of the reaction times for this task are likewise handled by Ibex farm, as are the stimuli for the acceptability judgment task.

4.5.3 Acceptability Judgment Task

The final experimental task is the acceptability judgment task, which is likewise accessed via a link, which participants are instructed to click following the self-paced reading task. At the beginning of the acceptability judgment task, the written instructions for the task appear on the screen. After reading the instructions, participants begin a short practice phase (10 example sentences) during which they receive feedback on their responses (no feedback is provided during the actual task). After the practice phase, they are given the option to begin the task or restart the instructions. A break is provided at the midpoint during the task.

The sentences for the self-paced reading task are presented in a 30 point traditional Arabic font, in a single line of text. Participants respond to each sentence using the Right or Left Control Key to indicate whether the sentence is acceptable (Right for 'Yes' and Left for 'No'). Following the participant's response, a new item will appear on the screen.

4.5.4 Vocabulary Survey

After the three experimental tasks, L2 participants are instructed to complete the Vocabulary Survey, during which they are asked to translate into English all the real Arabic words they responded to during the Lexical Decision task.

The overall duration of the experimental session was roughly 2 to 2.5 hours. When the experiment was completed, participants were debriefed (remote participants were instructed to email the primary investigator when they had completed the experimental procedure, at which point the primary investigator sent them the debriefing information).

All participants who completed the experimental procedure were compensated $40 for their time; remote participants were mailed a check after they mailed a signed receipt to the investigator.

5 Results

5.1 Lexical Decision Task

5.1.1 Data cleaning

The lexical decision task was administered to 33 L1 participants and 44 L2 participants. Participants whose accuracy on the lexical decision task fell below 70% were excluded from further analysis; thus 2 L1 participants and 6 L2 participants were excluded from analysis. (Post-experiment interviews revealed that the 2 L1 participants with less than 70% accuracy had skipped the instructions and misunderstood the LDT to be a matching task.) An additional 7 L2 participants were excluded from analysis because they did not complete the Vocabulary Survey after the LDT. This resulted in useable data from 31 L1 participants and 31 L2 participants.

An error in the coding of the data file for the experiment caused the timeout feature for the LDT to malfunction, such that when a participant waited longer than 3 seconds to respond, instead of timing out, the task remained on that trial until the participant pushed a response key. Thus, even though the instructions specified that participants should respond as quickly as possible while still being accurate, there was no timeout function to force them to hurry. Nevertheless, only 10.5% of the trials registered response latencies longer than 3000ms (which is where the timeout cutoff would have been). In order to approximate the timeout feature as nearly as possible, all trials with response times longer than 3 seconds were removed from further analysis.

The remaining data was cleaned by calculating each participant's mean response time (RT) and culling outliers that were further than 2.5 standard deviations from that mean.

Furthermore, only those L2 trials for which the L2 participant knew both the target and the prime word (as determined by the post-LDT vocabulary survey) were retained. Trials for which the participant failed to translate one or both of these words were discarded. (Between differences in accuracy between the language groups, and the vocabulary filtering just described, there are fewer L2 data points than L1 data points; specifically, there were an average of 64.6 useable trials (out of 80 possible trials) per L1 participant, and an average of only 34.1 useable trials per L2 participant.

The vast majority of excluded L2 data points (90.2% of them) were excluded because of a participant's failure to accurately define either the prime or the target word for that item on the vocabulary measure. L2 participants tended to outperform their conscious vocabulary knowledge, making accurate lexical decisions about words they could not translate; L2ers' average accuracy on the experimental task was 96% while their average accuracy on the Vocabulary Test was only 54%. This tendency for L2ers to outperform their vocabulary knowledge is unsurprising for several reasons; first, in order to perform accurately on the lexical decision task, a participant had only to recognize the target word, not the prime. For the Vocabulary Survey, however, a participant had to be able to define both. Second, L2 participants may have recognized some of the target words as Arabic words that they had seen or heard before, without remembering (or perhaps ever having known) their English translations. Finally, participants were excluded from the analysis if their accuracy was below 70%, but no lower cutoff was used to exclude participants who performed poorly on the vocabulary measure.

5.1.2 Response Time Summary

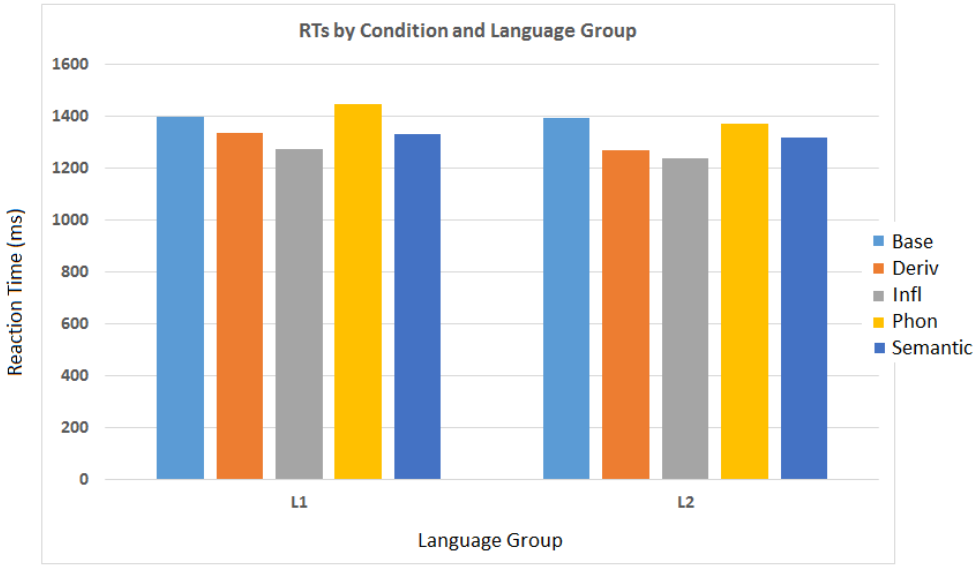
Table 5.1 shows the means and standard deviations for response time (RT) by condition for each language group.

Table 5.1 Response times by condition and language group

Language	Baseline		Derivational		Inflectional		Phonological		Semantic	
Group	Mean	StDev	Mean	StDev	Mean	StDev	Mean	StDev	Mean	StDev
L1	1398.8	394.3	1335.1	395.6	1273.4	347.8	1445.4	437.6	1331.1	394.3
L2	1390.0	395.0	1268.5	371.4	1235.6	339.1	1367.9	329.6	1316.4	347.0

The average mean RTs by language group across conditions are depicted graphically in Figure 5.1.

Figure 5.1 Response times by condition and language group



5.1.3 Effects of Conditions on Response Times

Four two-way ANOVAs were planned for the RT data, comparing each priming condition (Derivational, Inflectional, Phonological and Semantic) to the Baseline

condition. (In the by-subject analyses, the between-subjects factor was always Language Group (L1 vs. L2).)

Comparison of the Derivational condition to the Baseline showed a main effect of Derivational Relationship ($F1(1, 60) = 11.462$, $p < .001$; $F2(1,144) = 12.440$, $p = .001$) with no effect of Language Group ($F1(1,60) = .157$, $p = .693$; $F2(1,144) = 2.948$, $p = .088$) and no interaction between these factors ($F1(1,60) = 1.114$, $p = .295$; $F2(1,144) = 3.771$, $p = .055$).

Comparison of the Inflectional condition to the Baseline showed a main effect of Inflectional Relationship ($F1(1,60) = 27.791$, $p < .001$; $F2(1,144) = 18.239$, $p < .001$) with no effect of Language Group ($F1(1,60) = .067$, $p = .797$; $F2(1,144) = .018$, $p = .894$) and no interaction between these factors ($F1(1,60) = .299$, $p = .587$; $F2(1,144) = .127$, $p = .722$).

Comparison of the Phonological condition to the Baseline revealed no main effect of Phonological Relationship ($F1(1,60) = .166$, $p = .685$; $F2(1,144) = .123$, $p = .727$), nor any effect of Language Group ($F1(1,60) = .208$, $p = .650$; $F2(1,144) = .412$, $p = .522$) and no interaction between these factors ($F1(1,60) = 1.305$, $p = .258$; $F2(1,144) = .743$, $p = .390$).

Comparison of the Semantic condition to Baseline did show a main effect of Semantic Relationship ($F1(1,60) = 5.901$, $p = .018$; $F2(1,144) = 4.109$, $p = .047$) with no effect of Language Group ($F1(1,60) = .016$, $p = .899$; $F2(1,144) = .028$, $p = .866$) and no interaction between these factors ($F1(1,60) = .011$, $p = .919$; $F2(1,144) = .003$, $p = .956$).

Morphological vs. Phonological Effects

In order to compare effects of morphological priming to effects of phonological overlap, two additional two-way ANOVAs compared each morphological priming condition (Derivational and Inflectional) to the Phonological condition with the between-subjects factor Language Group (L1 vs. L2).

Comparison of the Derivational condition to the Phonological condition showed a main effect of Condition ($F1(1,60) = 20.473$, $p < .001$; $F2(1,144) = 13.736$, $p < .001$) with no effect of Language Group ($F1(1,60) = .573$, $p = .452$) in the by-subjects analysis (though there was an effect of language group in the by-items analysis ($F2(1,144) = 8.145$, $p = .005$)), and no interaction between these factors ($F1(1,60) = .056$, $p = .814$; $F2(1,144) = 1.358$, $p = .246$).

Comparison of the Inflectional condition to the Phonological condition likewise showed a main effect of Inflectional-Phonological Comparison ($F1(1,60) = 37.890$, $p < .001$; $F2(1,144) = 19.260$, $p < .001$) with no effect of Language Group ($F1(1,60) = .414$, $p = .523$; $F2(1,144) = 1.216$, $p = .272$) and no interaction between these factors ($F1(1,60) = .645$, $p = .425$; $F2(1,144) = .319$, $p = .573$).

To summarize the results of the lexical decision task, three of the priming conditions were significantly different from Baseline: the Derivational, Inflectional and Semantic priming conditions. Comparisons of each of these conditions to Baseline revealed no effect of, and no interaction with, Language Group.

Additionally, comparisons demonstrated each of the morphological priming conditions to be significantly different from the Phonological priming condition. These analyses also showed no effect of Language Group and no interaction between factors.

5.2 Self-Paced Reading

5.2.1 Data Cleaning

The self-paced reading task was administered to 31 L1 participants and 44 L2 participants. Participants whose accuracy on the task's comprehension questions fell below 70% were excluded from further analysis; thus 2 L1 participants and 5 L2 participants were excluded from analysis. The result was useable data from 29 L1 participants and 39 L2 participants. The remaining L1 participants' error rates ranged from 1% to 21%, with an average of 10%. The L2 participants' error rates ranged from 5% to 29%, with an average of 17%. These error rates suggest that participants in both language groups were reading the sentences for comprehension as they were instructed to, and generally understood the sentences[2].

The RT data was cleaned by first removing those trials with response times that were longer than 5 seconds. The remaining data was cleaned by calculating each participant's mean RT and removing outliers that were further than 2.5 standard deviations from that

[2] At least 3 L1 participants remarked on the difficulty of the SPR task during post-experiment interviews, noting that the length of the individual sentences combined with the length of the task made answering the comprehension questions surprisingly difficult. Nevertheless, it was apparent that these participants were trying and generally succeeding in following directions.

mean. These procedures cost 4.2% of the data. Each language group's mean RTs for each test position and each condition can be found in Table 5.2.

5.2.2 Response Time Summary

Table 5.2 Reading times by language group, condition and sentence region

	Base		Derivational		Inflectional		Semantic	
	Mean	St Dev	Mean	St Dev	Mean	St Dev	Mean	St Dev
L1								
-1	548.5	173.8	544.6	152.3	559.8	163.7	560.9	182.5
0	561.9	174.9	572.0	200.3	615.4	246.7	591.3	197.0
1	549.6	146.5	628.3	198.1	632.7	190.3	684.8	262.7
2	522.9	158.7	558.6	148.0	551.9	184.1	588.5	189.6
3	497.8	177.3	536.1	178.1	503.2	175.8	555.6	206.6
L2								
-1	1095.1	490.7	1081.3	460.5	1084.9	427.1	1069.4	469.5
0	1155.4	508.2	1250.8	576.5	1404.3	580.7	1242.6	505.6
1	983.0	333.5	980.4	342.7	1058.7	380.3	1072.9	365.0
2	848.2	369.2	842.3	264.2	917.1	393.7	834.8	243.7
3	847.1	347.8	849.7	312.3	855.8	313.2	908.0	319.6

Figure 5.2 depicts mean RTs by position and condition in line graph form for the L1 group, while Figure 5.3 does the same for the L2 group.

Figure 5.2 L1 Reading times by sentence position and condition

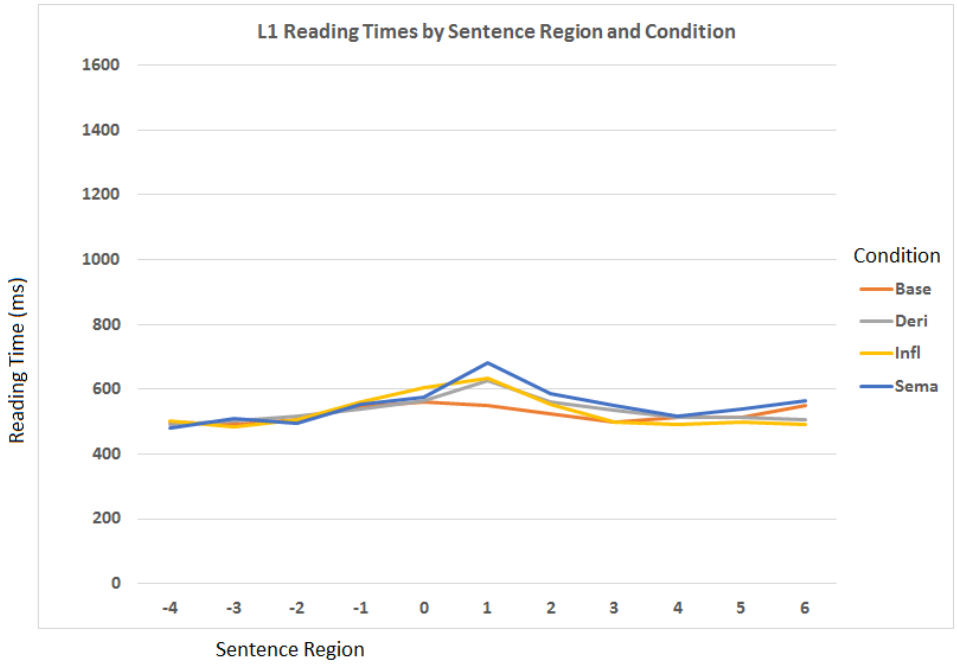
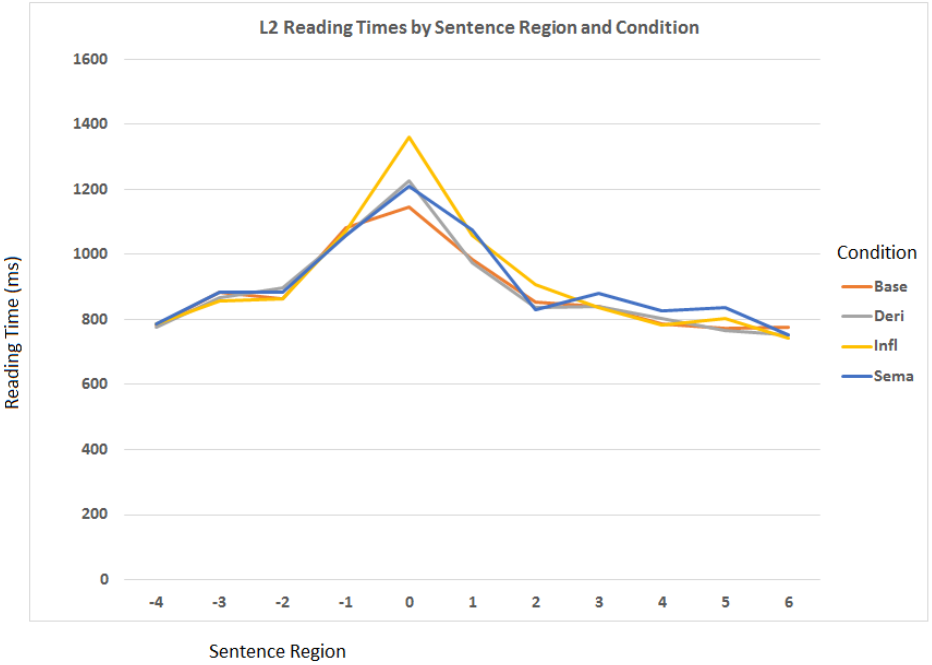


Figure 5.3 L2 Reading times by sentence position and condition



5.2.2 Effects of Conditions on Reading Times

To investigate the effects of each error condition on reading times, four sets (one for each sentence region) of three 2x2 ANOVAs were planned, one comparing each error condition (Derivational, Inflectional and Semantic) to the Baseline condition. (In the by-subjects analysis, Language Group (L1 vs. L2) was always the between-subjects factor. By items analyses were also conducted, wherein the between- and within- factors were reversed.)

Comparison of the Derivational condition to the Baseline revealed no effect of Condition (*F1(*1, 66) = .175, $p = .677$; *F2*(1, 118) = 2.335, $p = .129$) in the precritical region (position -1). There was a significant effect of Language Group (*F1(*1, 66) = 36.326, $p < .001$; *F2*(1, 118) = 381.115, $p < .001$) in this region, but no interaction between factors (*F1(*1, 66) = .056, $p = .814$; *F2*(1, 118) = .018, $p = .893$). The lack of any effect of condition in the precritical region is to be expected, as all the conditions are identical before they reach the critical region. In the critical region (position 0), the effect of Condition becomes significant in the by-subjects analysis, though not in the by-items analysis (*F1(*1, 66) = 8.542, $p = .005$; *F2*(1, 118) = 2.335, $p = .129$). The effect of Language group is significant here as well (*F1(*1, 66) = 37.480, $p < .001$; *F2*(1, 118) = 465.776, $p < .001$), and the interaction between Condition and Language Group is significant in by-subjects but not in by-items analysis(*F1(*1, 66) = 5.569, $p = .021$; *F2*(1, 118) = 1.821, $p = .180$). Simple comparisons revealed a significant effect of Condition in the L2 group in by-subject analysis but not in by-items analysis (*t1*(38) = -3.198, p = .003; *t2*(118) = -1.523, $p = .131$) and no condition effect in the L1 group (*t1*(28) = -.859, p =.397; *t2*(118) = .728, $p = .468$). In the first spillover region (position 1), the effect of

Condition remains significant in the by-subjects but not the by-items analysis (*F1(*1, 66) = 5.081, *p* = .028; *F2*(1, 118) = 1.443, *p* = .232). Effects of language group are significant in both analyses (*F1(*1, 66) = 34.684, *p* < .001; *F2*(1, 118) = 245.602, *p* < .001) and the interaction between those factors is significant by-subjects but not by-items (*F1(*1, 66) = 5.817, *p* = .019; *F2*(1, 118) = 2.768, *p* = .099). Simple comparisons now reveal a significant effect of Condition in the L1 group (*t1*(28) = -4.235, *p* < .001; *t2*(118) = -3.892, *p* < .001), but not the L2 group (*t1*(38) = .104, *p* = .918; *t2*(118) = .056, *p* = .995), which is the reverse of what was just seen in the critical region. By the second spillover region (position 2), there is no longer any significant effect of Condition (*F1(*1, 66) = .356, *p* = .553; *F2*(1, 118) = .131, *p* = .718). The only significant effect here is of Language Group (*F1(*1, 66) = 26.144, *p* < .001; *F2*(1, 118) = 169.070, *p* < .001), and there is no interaction between these factors (*F1(*1, 66) = .696, *p* = .407; *F2*(1, 118) = 1.780, *p* =.185).

Comparison of the Inflectional Error condition to the Baseline revealed no effect of Condition (*F1(*1, 66) = .001, *p* = .976; *F2*(1, 118) = .034, *p* = .854) in the precritical region (position -1). Here the only significant effect is of Language Group (*F1(*1, 66) = 37.267, *p* < .001; *F2*(1, 118) = 398.954, *p* < .001), and there is no interaction between these factors (*F1(*1, 66) = .001, *p* = .976; *F2*(1, 118) = .213, *p* = .645). In the critical region, effect of Condition becomes significant (*F1(*1, 66) = 39.292, *p* < .001; *F2*(1, 118) = 14.739, *p* < .001). There are likewise significant effects of Language Group (*F1(*1, 66) = 43.864, *p* < .001; *F2*(1, 118) = 530.706, *p* < .001) and interaction between Condition and Language Group (*F1(*1, 66) = 16.385, *p* < .001; *F2*(1, 118) = 8.240, *p* = .005). Simple comparisons indicate significant effects of Condition in both the L1 group (*t1*(28)

= -2.720, p = .011; *t2*(118) = -2.314, p = .022), and the L2 group (*t1*(38) = -6.401, p < .001; *t2*(118) = -3.561, p = .001). In the first spillover region, effects of Condition (*F1(*1, 66) = 19.427, p < .001; *F2*(1, 118) = 4.942, p = .028) and Language Group (*F1(*1, 66) = 38.160, p = .001; *F2*(1, 118) = 246.056, p < .001) remain significant, but there is no longer any interaction between them (*F1(*1, 66) = .042, p = .839; *F2*(1, 118) = .062, p = .804). In the second spillover region, effects of Condition are significant in by-subject but not in by-items analyses (*F1(*1, 66) = 6.549, p = .013; *F2*(1, 118) = 1.934, p = .167). Language Group effects are significant (*F1(*1, 66) = 21.948, p < 001; *F2*(1, 118) = 167.147, p < .001). There is no interaction between them (*F1(*1, 66) = 1.090, p = .300; *F2*(1, 118) = .067, p = .796) in this region either.

Comparison of the Semantic Error condition to the baseline revealed no effect of Condition (*F1(*1, 66) = .086, p = .771; *F2*(1, 118) = .028, p = .866) in the precritical region. The only significant effect in the precritical region is of Language Group (*F1(*1, 66) = 33.621, p < .001; *F2*(1, 118) = 362.082, p < .001), and there is no interaction between Language Group and Condition (*F1(*1, 66) = .708, p = .403; *F2*(1, 118) = .256, p = .614). In the critical region, the effect of Condition becomes significant in by-subjects but not by-items analysis (*F1(*1, 66) = 6.873, p = .011; *F2*(1, 118) = 2.312, p = .131). Language Group is also significant here (*F1(*1, 66) = 41.719, p < .001; *F2*(1, 118) = 383.434, p < .001), but there is no interaction between these factors (*F1(*1, 66) = 1.689, p = .198; *F2*(1, 118) = 1.639, p = .203). In the first spillover region, the effect of Condition is significant (*F1(*1, 66) = 18.392, p < .001; *F2*(1, 118) = 8.047, p = .005), as does the effect of Language Group (*F1(*1, 66) = 35.927, p < .001; *F2*(1, 118) = 268.411, p < .001), and there is still no interaction between them (*F1(*1, 66) = .746, p = .391; *F2*(1, 118) =

1.381, p = .242). By the second spillover region, there is no longer any effect of Condition (*F1(*1, 66) = 1.329, p = .253; *F2*(1, 118) = .465, p = .497). The only significant effect here is of Language Group (*F1(*1, 66) = 22.339, p < .001; *F2*(1, 118) = 151.177, p < .001); there is no interaction between these factors (*F1(*1, 66) = 3.054, p = .085; *F2*(1, 118) = 2.411, p = .123).

For ease of reference, Tables 5.3 and 5.4 display the simple comparison outcomes (by subjects and by item analyses, respectively) by language group, condition and region of interest.

Table 5.3 Simple comparison significance by group, condition and region (by subjects)

			sentence region							
			precritical (-1)		critical (0)		spillover 1		spillover 2	
L Group	Condition	df	t	p	t	p	t	p	t	p
L1	Deriv	28	0.255	0.801	0.859	0.397	**-4.235**	**< 0.001**	**-3.639**	**0.001**
	Infl	28	-0.912	0.37	**2.72**	**0.011**	**-4.541**	**< 0.001**	-1.453	0.157
	Sem	28	0.884	0.384	1.755	0.09	**-5.032**	**< 0.001**	**-4.813**	**< 0.001**
L2	Deriv	38	0.4	0.691	**-3.198**	**0.003**	0.104	0.918	0.14	0.889
	Infl	38	0.332	0.742	**-6.401**	**< 0.001**	**-2.714**	**0.01**	**-2.342**	**0.025**
	Sem	38	0.684	0.493	**-2.408**	**0.021**	**-2.216**	**0.033**	0.358	0.722

Table 5.4 Simple comparison significance by group, condition and region (by items)

			sentence region							
			precritical (-1)		critical (0)		spillover 1		spillover 2	
L Group	Condition	df	t	p	t	p	t	p	t	p
L1	Deriv	118	0.238	0.812	-0.728	0.468	**-3.892**	**< .001**	**-2.395**	**0.018**
	Infl	118	-0.789	0.431	**-2.314**	**0.022**	**-4.248**	**< .001**	**-2.056**	**0.042**
	Sem	118	-0.383	0.702	-0.737	0.463	**-5.443**	**<.001**	**-3.153**	**0.002**
L2	Deriv	118	0.217	0.829	-1.523	0.131	0.056	0.955	0.384	0.702
	Infl	118	0.103	0.918	**-3.561**	**0.001**	-1.165	0.247	-0.945	0.346
	Sem	118	0.344	0.731	-1.471	0.144	-1.268	0.207	0.265	0.791

5.2.3 Summarizing Self-Paced Reading Results

To summarize, across regions and conditions, main effects of Language Group were always significant. From Figures 5.2 and 5.3 it is apparent that L1 reading times on the

whole were faster than L2 reading times, hence the Language Group effects across the board.

In the precritical region (position -1), no error condition effects were significant, which makes sense because in this region participants have not yet encountered the error word.

By the critical region (position 0), L2 participants' RTs are significantly different from baseline in all three error conditions, while L1 participants' RTs are only significantly different from baseline in the Inflectional condition.

By the first spillover region (position 1), L1 participants' RTs are now significantly different from baseline in all three error conditions, whereas L2 participants' RTs are significantly different from baseline in the Inflectional and Semantic conditions only.

By the second spillover region (position 2), ANOVAs show only significant effects of Inflectional Error and of Language Group but no interaction between them.

5.3 Acceptability Judgment Task

5.3.1 Data cleaning

The acceptability judgment task was administered to 31 L1 participants and 44 L2 participants. Participants were excluded from further analysis for this task if their accuracy was below 50%. In this way, 1 L1 participant and 6 L2 participants were excluded. The result was useable data from 30 L1 participants and 38 L2 participants.

5.3.2 Accuracy Score Summary

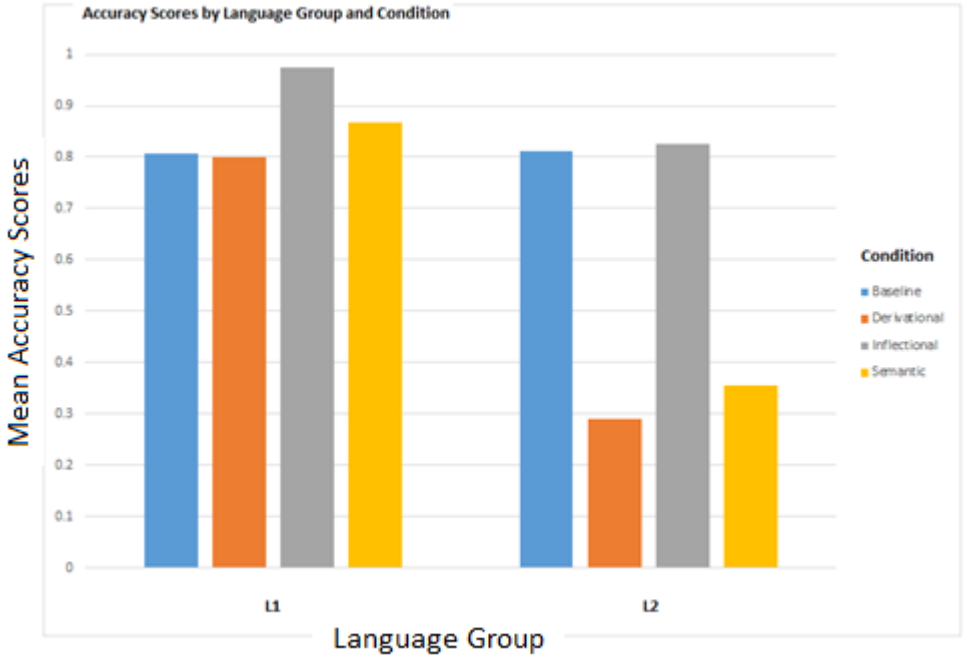
Table 5.5 shows the means and standard deviations for accuracy scores by condition for each language group.

Table 5.5 Accuracy scores by language group and condition

	Baseline		Derivational		Inflectional		Semantic	
	Mean	StDev	Mean	StDev	Mean	StDev	Mean	StDev
L1	0.81	0.15	0.80	0.18	0.97	0.05	0.87	0.18
L2	0.81	0.11	0.29	0.17	0.83	0.22	0.35	0.24

The average mean accuracy scores by language group across conditions are depicted graphically in Figure 5.4.

Figure 5.4 Mean accuracy scores by language group and condition



5.3.3 Effects of Conditions on Accuracy Scores

To compare the effects of the conditions on accuracy scores, three 2X2 ANOVAs were planned comparing each of the error conditions (Derivational, Inflectional and Semantic) to the Baseline condition, with Language Group (L1 vs. L2) as the between-subjects factor in all three.

Comparison of the derivational error condition to the baseline revealed a significant effect of Condition ($F1(1, 66) = 78.441$, $p < .001$; $F2(1, 118) = 117.616$, $p < .001$), as well as a significant effect of Language Group ($F1(1, 66) = 113.300$, $p < .001$; $F2(1, 118) = 114.692$, $p < .001$), and a significant interaction between them ($F1(1, 66) = 74.540$, $p < .001$; $F2(1, 118) = 111.296$, $p < .001$). Simple comparisons showed a significant effect of Condition in the L2 group only ($t1(37) = 15.307$, $p < .001$; $t2(118) = 17.840$, $p < .001$), in the L1 group this contrast was not significant ($t1(29) = .129$, $p = .898$; $t2(118) = .579$, $p = .564$).

Comparison of the inflectional error condition to the baseline showed a significant main effect of Condition ($F1(1, 66) = 15.460$, $p < .001$; $F2(1, 118) = 24.387$, $p < .001$), as well as a significant effect of Language Group ($F1(1, 66)= 6.257$, $p = .015$; $F2(1, 118) = 19.034$, $p < .001$), and a significant interaction between them ($F1(1, 66) = 11.030$, $p = .001$; $F2(1, 118) = 17.207$, $p < .001$). Simple comparisons indicated a significant effect of Condition in the L1 group ($t1(29) = -5.301$, $p < .001$; $t2(118) = -5.999$, $p < .001$) but not in the L2 group ($t1(37) = -.433$, $p = .668$; $t2(118) = -.521$, $p = .603$).

Comparison of the semantic error condition to the baseline showed a significant main effect of Semantic Error ($F1(1, 66) = 34.603$, $p < .001$; $F2(1, 118) = 81.712$, $p < .001$), as

well as a significant effect of Language Group (*F1(*1, 66) = 86.575, $p < .001$; *F2*(1, 118) = 156.758, $p < .001$), and a significant interaction between them (*F1(*1, 66) = 58.623, $p < .001$; *F2*(1, 118) = 152.283, $p < .001$). Simple comparisons revealed a significant effect of Condition in the L2 group (*t1(*37) = 10.178, $p < .001$; *t2*(118) = 16.898, $p < .001$) but not in the L1 group (*t1(*29) = -1.189, $p = .244$; *t2*(118) = -1.776, $p = .078$).

To summarize the results of the AJT task, it is apparent that L2 learners are driving the effects of condition in both the Derivational and Semantic error conditions. Figure 5.4 shows that L2 participants' accuracy scores in these two conditions are well below 50%, that is to say, they are performing at below chance accuracy. Conversely, it is L1 participants who are driving the effect of condition in the Inflectional error condition. In Figure 5.4 it is apparent that L1 participants' accuracy in the Inflectional condition is higher than it is in the Baseline condition; indeed the Inflectional condition has the highest mean accuracy of any condition. This pattern points to a difference between the Inflectional error condition and the other two error conditions. The implication is that the inflectional errors were easier to make judgments about than the other two error conditions. The next section delves further into the reasons for the patterns observed across the three experimental tasks and their implications for the theoretical claims described in section 3.

6 Discussion and Conclusions

This section revisits the results just described and considers their implications for theoretical approaches to L2 morphological processing.

6.1 LDT Discussion

6.1.1 L1 LDT Findings

Analyses of the lexical decision task data demonstrated significant effects of morphological priming in both the derivational and inflectional conditions relative to baseline. Furthermore, both kinds of morphological priming were distinct from effects of phonological overlap, in that response times in both morphologically related conditions were significantly faster than those in the phonological condition. As Figure 5.1 shows, mean RTs in the phonological condition were actually slower than in the baseline condition for the L1 group, suggesting that the effects of phonological overlap alone trend towards inhibition for native Arabic speakers. Thus, priming in the morphologically related conditions cannot be explained in terms of sheer phonological overlap.

Significant effects of semantic priming were also observed relative to baseline.[3] This pattern of priming in the derivationally related conditions accords with previous research that has found derivational priming in native speakers of English (Marslen-Wilson, Tyler, Waksler & Older, 1994), Hebrew (Bentin and Feldman, 1990; Frost, Forster and Deutsch, 1997), German (Neubauer & Clahsen, 2010), and Polish (Reid and Marslen-Wilson, 2000). That L1 participants should also show priming in the inflectional condition fits

[3] While Arabic morphology allows for disambiguation between morphological and semantic priming effects, the current study was not designed to investigate this difference. Boudelaa and Marslen-Wilson (2000) showed, however, that morphological priming effects obtain in native speakers of Arabic even in the absence of semantic relatedness, just as Bentin and Feldman (1990) found for native speakers of Hebrew.

with previous research that demonstrated inflectional priming in native speakers of Dutch (Drews & Zwitserlood, 1995), German (Sonnenstuhl, Eisenbeiss & Clahsen, 1999), English (Marslen-Wilson, Hare & Older, 1993), and Italian (Orsolini & Marslen-Wilson, 1997).

6.1.2 L2 LDT Findings

The lack of any effect of language group, or any interaction with language group in comparisons between the priming conditions and the baseline condition suggest that mean RT patterns in the L2 group were not significantly different from those in the L1 group. This conclusion is supported by simple comparisons between conditions in the L2 data, which revealed significant effects of derivational as well as inflectional priming compared to baseline. Simple comparisons also showed that the two morphological priming conditions differed significantly from the phonological condition. As can be seen in Figure 5.1, L2 participants did not show the trend towards phonological inhibition that L1 participants showed. Nevertheless, the phonological condition was not significantly different from baseline for the L2 participants.

This pattern of morphological priming in the derivational condition fits with previous research demonstrating derivational morphological priming in L2 learners of English (Feldman, Kostic, Basnight-Brown, Durdevic & Pastizzo, 2010; Diependaele et al., 2011), German (Silva & Clahsen, 2008), Turkish (Kirkici & Clahsen, 2013) and Arabic (Freynik, Gor & O'Rourke, submitted). That L2 participants should also show priming in the inflectional condition fits with previous research that found inflectional priming in L2 learners of English (Basnight-Brown, Chen, Hua, Kostic and Feldman, 2007), Swedish (Portin, Lehtonen, Harrer, Wande, Niemi & Laine, 2008), and Russian (Gor & Cook,

2010), and supports the conclusion that L2 acquisition of Arabic is not unlike L2 acquisition of other languages in this respect. Simple comparisons among conditions in the L2 data also revealed that semantic priming was not significantly different from baseline for L2 participants, though they trended towards faster RTs in that condition. This finding corresponds to earlier research that found that L2 learners of Arabic did not show significant semantic priming (Freynik, Gor & O'Rourke, submitted). Such results suggest that L2 participants' semantic processing is deficient in some way. It is possible that the semantic information accompanying L2 learners' Arabic lexical entries is underspecified, or that connections between lexical entries are weaker for an L2 learner than they are for a native speaker. This idea will be revisited during the discussion of results from the acceptability judgment task below.

6.1.3 Comparing L1 and L2 LDT Findings

On the whole, L1 and L2 participants showed similar priming patterns among conditions. This fits with other research comparing L1 and L2 performance in an Arabic lexical decision task with morphological priming (Freynik, Gor & O'Rourke, submitted). Such findings suggest that both L1 and L2 learners of Arabic make use of morphological structure during lexical access, even when that morphology is discontinuous, as it is in the derivational condition. A difference between L1 and L2 RT patterns concerns semantic priming effects; while L2 participants show a trend towards semantic priming, the effect is not significant for them the way it is for L1 participants. Thus, while L2 learners seem relatively similar to native speakers in their use of morphological information, they appear deficient in their use of semantic information.

One unusual aspect of these results is the relative comparability of mean RTs between language groups. Typically L1 participants have faster RTs than L2 participants; for example, while Freynik, Gor & O'Rourke (submitted) observed the same general pattern of priming in L2 learners as in L1 learners, the L2 learners in their study had slower RTs overall, resulting in an effect of Language group but no interaction between language group and condition. Conversely, the current study found no significant difference between the response times of the two language groups. One possible explanation for this difference between the two studies concerns the fact that, while roughly half the data for Freynik, Gor & O'Rourke (submitted) was collected locally in the lab on campus, all of the data for the current study was collected remotely, and many of the L1 participants were living abroad. It is possible that this population of L1 participants were less experienced with button-press reaction time experiments than their L2 counterparts, for while recruitment efforts began with university listserv emails, after debriefing each participant was encouraged to pass flyers advertising the experiment on to acquaintances that might be interested. In addition to experiment-participation experience, testing conditions may have differed between the two language groups (internet cafes, for example, are no longer common in the United States but remain common in Cairo and Beirut), and the lack of a timeout feature may have exacerbated any tendency towards slower RTs that these differences in experience levels and testing conditions may have created. Future remote data collection should include more detailed inquiry into testing conditions as well as participants' prior experiences with reaction-time-based experiments. Nevertheless, after trimming data from trials with RTs longer than the intended 3000ms timeout mark (which was 10.5% of all data), the priming patterns in the

remaining L1 data conformed to the patterns observed by Boudelaa & Marslen-Wilson (2000; 2001; 2004; 2011), and the L2 RTs conformed to L2 Arabic priming patterns previously observed (Freynik, Gor & O'Rourke, submitted).

6.2 Self-Paced Reading Task Discussion

6.2.1 L1 Self-Paced Reading Findings

Analyses of the self-paced reading task indicated significant effects of all three error conditions on L1 participants' reading times. L1 participants responded to the inflectional error condition with slower reading times in the critical region. Slowdowns in the Inflectional error condition were expected for L1 participants, and accord with previous research that found slowdowns in response to inflectional errors in native speakers of English (Pearlmutter, Garnsey & Bock, 1999; Jiang, 2007), Spanish (Sagarra & Herschensohn, 2010) and Hebrew (Deutsch, 1998).

The effects of the derivational error condition on L1 participants' RTs were not apparent until the first spillover region, but here they were significant. Such slowdowns in response to derivational errors accord with results like those of Clahsen and Ikemoto (2012) who found that native speakers of Japanese responded to sententially-inapprorpriate nominalization morphemes with slower reading times. Thus, L1 participants' behavior in the two morphological error conditions indicates that native speakers of Arabic respond online to morphological errors in the same way as native speakers of other languages.

Like the derivational error condition, the semantic error condition affected L1 participants' reading times in the first spillover region. Semantic anomalies have been

shown to result in slower reading times for native speakers of Italian (Vincenzi, 2003). Interestingly, Vincenzi (2003) compared the effects of semantic anomalies to the effects of syntactic anomalies on Italian L1 participants' reading times and found that the slowdowns resulting from semantic anomalies emerged later than the slowdowns resulting from syntactic ones. Arabic L1 participants in the current study appear to evince the same pattern in the semantic condition. The fact that their responses to derivational errors were also slower than their responses to inflectional errors suggests that derivational errors may be processed in a way similar to semantic errors, but this conjecture is merely speculative at this point.

6.2.2 L2 Self-Paced Reading Findings

Analyses of L2 participants' reading times indicated significant effects of all three error conditions in the critical region. That L2 learners should demonstrate online sensitivity to inflectional violations was not a given. On the one hand some previous research provides evidence of online sensitivity to inflectional errors in L2 learners. Hopp (2006) and Jackson (2008) found that proficient L2 learners of German showed slower reading times in response to pragmatically unexpected case-marking. Similarly, Sagarra and Herschensohn (2010) found that proficient L2 learners of Spanish responded to number and gender agreement errors with slower reading times. On the other hand, however, Jiang (2007) found that Chinese learners of English showed no online slowdowns in response to number agreement violations. The current results accord with those of Hopp (2006), Jackson (2008), and Sagarra and Herschensohn (2010), and conflict with those of Jiang (2007), but it's important to note that the English number agreement violations examined by Jiang (2007) differ from the Arabic number agreement violations examined

in the current study in a number of important ways, including their salience and their familiarity to the learners in question. This contrast will be addressed in more detail during the discussion of theoretical approaches to L2 morphological deficiency below.

The derivational error condition also significantly affected L2 learners' reading times in the critical region. That L2 learners should be sensitive to violations of Arabic pattern morphology is somewhat surprising. While the lexical decision tasks in the current study and in Freynik, Gor & O'Rourke (submitted) demonstrated L2 learners' ability to make use of Arabic derivational morphemes during lexical access, the mere recognizing of a word could still be achieved via comparably shallow processing. (Indeed, the discrepancies between L2 learners' accuracy on the LDT task and their ability to define those same words post-task indicated that they were able to recognize Arabic words that they could not define.) That L2 learners appear to notice derivational violations in sentential contexts suggests that they have access to the features the derivational morphemes in question encode. Further, L2 learners' sensitivity to such features in an online measure implies automatized, (as opposed to offline, declarative) knowledge. In Jiang's words, "It is believed that such sensitivity [as indicated by a delay while reading incorrect sentences] can be observed only when the involved L2 knowledge has been highly integrated and is automatically available."

Like the other two error conditions, the semantic error condition significantly affected L2 learners' reading times in the critical region. Slowdowns in the semantic condition were expected in both language groups, and echoed findings like those of Vincenzi (2003).

An alternate explanation for L2 slowdowns in the derivational and semantic conditions could be that L2 participants were simply less familiar with the verb forms in these two conditions than they were with the verb forms in the baseline condition. A few factors weigh against this conclusion, however. First, the majority (roughly ¾) of the derived verb pairs used in the baseline and derivational conditions were taken from A Frequency Dictionary of Arabic: Core Vocabulary for Learners (Buckwalter & Parkinson, 2010) which lists the 5000 most frequent words in modern usage in the Arabic language. And nearly all of the verbs used in the semantic condition were taken from this dictionary. That is, the verbs in the baseline, derivational and semantic conditions alike were comparably frequent verb forms that learners were expected to be familiar with. Furthermore, in many of the trials, the verb in the derivational error condition was actually the Form 1 verb, while the verb in the baseline condition was another form. That is to say, the verb in the baseline condition was baseline because it was grammatically and semantically appropriate, not because it was a more "basic" derivation of the verbal root.

6.2.3 Comparing L1 and L2 Self-Paced Reading Findings

Analyses of self-paced reading data indicated a significant effect of language group. From figures 5.2 and 5.3 it is apparent that L1 reading times were significantly faster than L2 reading times.

Both L1 and L2 participants responded to all three error conditions (inflectional, derivational and semantic) with significantly slower reading times between the critical and spillover regions, but the breakdown by condition and region differed for the two language groups. That is, L1 participants showed immediate slowdowns in the critical

region in response to inflectional errors, but effects of derivational and semantic anomalies did not significantly affect L1 reading times until the first spillover region. As mentioned above, this pattern corresponds to one observed by Vincenzi (3003), wherein native speakers of Italian evinced slowdowns more quickly in response to syntactic anomalies than to semantic anomalies.

Conversely, L2 reading times showed significant effects of all three error conditions in the critical region. This difference between the language groups is comparably unimportant compared to the fact that both groups evinced slowdowns at all, and did so for all the error conditions. Here it is also important to remember, however, that the mere fact of similar slowdowns in both language groups does not entail that similar underlying processes were going on in both L1 and L2 participants' minds. This caveat will be expanded during the discussion of research questions below.

6.3 Acceptability Judgment Task Discussion

6.3.1 L1 Acceptability Judgment Task Findings

Among L1 participants, average accuracy scores were above 80% in all conditions of the acceptability judgment task. The only L1 condition that differed significantly from baseline (or indeed from any of the others) was the inflectional error condition (97% vs. 81% accuracy in the Baseline condition). This distinction can be understood in terms of the categorical wrongness of the violations in the Inflectional condition compared to those of the other conditions. That L1 learners of Arabic should make accurate judgments about inflectional errors conforms to previous research that found accurate judgments

about inflectional errors in native speakers of Spanish (Sagarra & Herschensohn, 2010), German (Clahsen, Sonnenstuhl & Blevins, 2003), and English (Jiang, 2004; 2007).

As mentioned above in section 4, a sentence like, "The monument builds…" (inflectional violation) is grammatically wrong, and differs from the semantic wrongness of "The monument are built… " (derivational violation in Arabic) or even "The monument regrets…" (semantic violation). While L1 participants were generally willing to reject sentences in the latter two conditions, they did so less categorically than they did in the inflectional condition. Some L1 participants mentioned, during post-experiment interviews, that they had judged some sentences acceptable though they could only imagine them happening in storybook contexts. This was despite the fact that examples of unacceptable sentences presented during pre-task instructions included semantic violations like, "He drinks the coffee with sugar and socks," followed by the explanation, "This sentence is bad because it doesn't make sense."

Nevertheless, L1 participants were relatively accurate in identifying derivational errors, a result that conforms to Clahsen and Ikemoto's (2012) finding that native speakers' acceptability judgments distinguished between appropriate and inappropriate nominalization morphemes in Japanese.

L1 participants' less-than-ceiling accuracy in the Baseline condition likewise merits explanation. Though experimental items were vetted by three different native speakers and those in the baseline condition were deemed acceptable, later L1 participants in post-experiment interviews mentioned rejecting sentences due to problems with the particles that accompanied some of the verbs (as well as due to disagreements regarding content

such as the probable cost of soup). Though Modern Standard Arabic is supposedly a standardized language understood throughout the Arabic-speaking world, there are preferences regarding use of prepositions and particles that differ from region to region, and as participants in the current study hailed from 10 different countries, there was some unavoidable variance in judgments about the acceptability of some items.

6.3.2 L2 Acceptability Judgment Task Findings

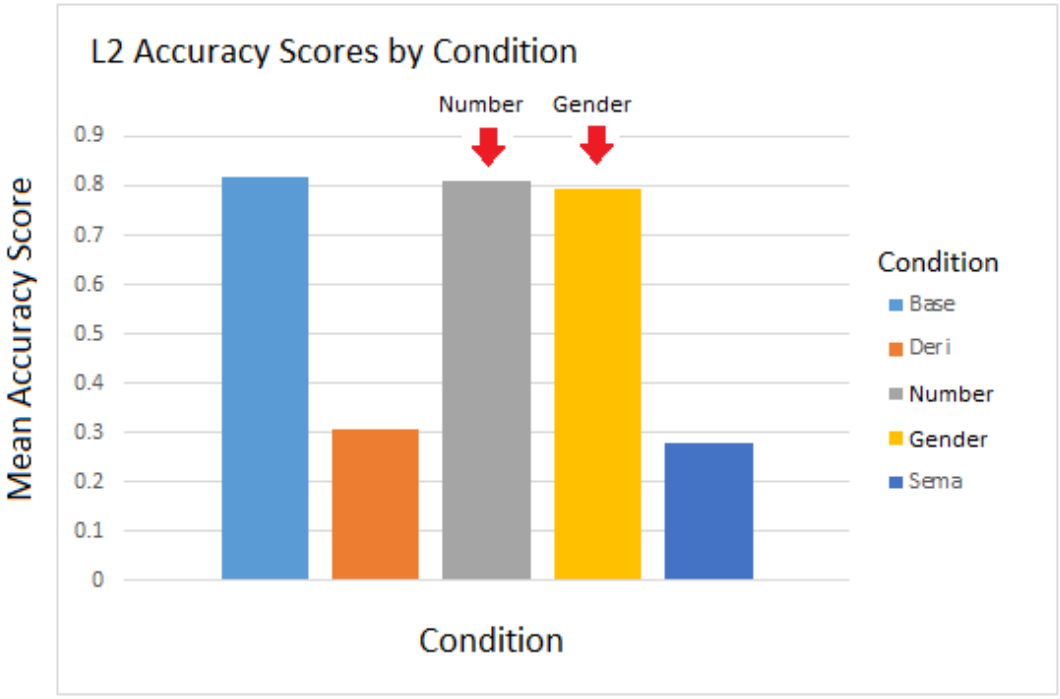
As mentioned above, condition played a greater role in predicting L2 participants' accuracy than it did L1 participants'. While L2 participants scored 81% accuracy for baseline and 83% accuracy for inflectional violations, they scored well below chance in the derivational and semantic conditions (29% and 35% respectively). As with the L1 participants, L2 participants' accuracy in the inflectional condition can be understood in terms of the kind of error that appeared in it: categorically wrong and easy to point to. That L2 learners should make comparably accurate judgments about inflectional errors in an AJT task fits previous research like Sagarra and Herschensohn (2010) who found that L2 learners of Spanish made accurate judgments about gender and number agreement errors.

More broadly, L2 participants' behavior across conditions in the current study is best accounted for by a strong acceptance bias, one that is only overcome in instances where a clear, grammatical error can be identified. Such an acceptance bias explains not only L2 participants' relative accuracy in the baseline condition (where an acceptance judgment was always correct) but their below-chance performance in the derivational and semantic conditions. Whereas random behavior would have led to chance performance in these

conditions, L2 participants performed below chance because their tendency was to judge these (incorrect) sentences acceptable.

That grammatical vs. lexico-semantic acceptability was the relevant predictor of L2 participants' behavior in the AJT was underscored by participants' performance in filler condition sentences. One set of filler sentences manipulated subject-verb gender agreement. L2 participants' accuracy in this condition (78%) was comparable to their accuracy (83%) in the inflectional condition (which conversely always involved subject-verb number agreement). By looking only at the critical items, one might conclude that L2 learners' accuracy in the inflectional (number agreement) condition had to do with L1 familiarity, as subject-verb number agreement is a feature that English and Arabic share. However, Arabic subject-verb gender agreement has no English parallel, and L2 learners still made accurate judgments about it.

Figure 6.1 L2 AJT Accuracy by Condition (including gender agreement)



6.3.3 Comparing L1 and L2 Acceptability Judgment Task Findings

Both L1 and L2 participants were most accurate in the inflectional error condition, implying that this condition was easier to provide judgments about than the other conditions. This relative facility is probably due to the inflectional errors being obvious and categorical in a way that the derivational and semantic anomalies were not. L1 and L2 participants also exhibited comparable accuracies in the baseline condition, but this appears to be coincidental. L1 participants should have been at ceiling accuracy in their native language, but variance in regional preferences regarding particle and preposition choices seems to have dragged their acceptance levels down in this condition. By contrast, L2 participants' acceptance bias worked in their favor in the baseline condition. That the two groups' accuracy scores converged in this condition seems to be an example of different underlying processes giving rise to a similar-looking surface manifestation, and should serve as a reminder for caution in the interpretation of such results across the board.

The greatest differences between L1 and L2 participants were in the derivational and semantic conditions. L1 participants' accuracies in these conditions were not significantly different from baseline, whereas L2 participants' accuracies in these conditions were well below chance. When L2 participants' low accuracies in these conditions is compared to their ~80% accuracies in the inflectional (number agreement) error condition and the gender agreement error condition, it seems that the relevant distinction among conditions for L2 participants is lexico-semantic vs. syntactic error. What the semantic and derivational conditions have in common is the lexico-semantic nature of the error. And

L2 participants were less willing to judge a sentence unacceptable on lexico-semantic grounds than on strictly grammatical grounds.

In this way the acceptability judgment data provides a telling contrast to the self-paced reading data. L2 participants in the self-paced reading task seemed to demonstrate sensitivity to derivational, inflectional and semantic violations alike in their online reading times, suggesting that at some level they registered that something was amiss in all the error conditions. However, when the task called for a conscious, metalinguistic judgment (the subject-verb pairs for the critical items in the SPR and AJT sentences were the same), L2 participants were not confident enough to reject sentences with derivational or semantic violations. This pattern of results is counter-intuitive; the more common result is for L2 participants to demonstrate conscious, metalinguistic awareness of a linguistic phenomenon that they do not yet have automated control of. However, the opposite pattern has also been reported in studies like Tokowicz & Macwhinney (2005) and McLaughlin, Osterhout & Kim (2004) that found L2 learners seemingly outperforming their conscious, declarative knowledge when online measures like self-paced reading times and event related potentials were used.

6.4 Discussion of Research Questions

As laid out in section 3, the research questions that the current study aimed to address were the following:

1. Are L2 learners more sensitive to Arabic derivational morphology than to Arabic inflectional morphology at the lexical level?

2. Is their morphological sensitivity limited to the lexical level or can L2 learners make use of this morphological information during sentence processing (i.e., how "deep" and automatic is their knowledge of this morphology)?
3. How does automatic or integrated L2 knowledge of morphology compare to explicit, conscious L2 knowledge of morphology?

6.4.1 Research Question 1

Regarding research question 1 (whether L2 learners are more sensitive to Arabic derivational morphology than to Arabic inflectional morphology at the lexical level), the results of the lexical decision task answer in the negative. L2 learners appear to be equally sensitive to both Arabic derivational and inflectional morphology at the lexical level. As described above, analyses of the lexical decision task's response time data demonstrated significant effects of morphological priming in both the derivational and inflectional conditions relative to baseline. Effects of derivational and inflectional morphological priming were also significantly different from the phonological condition. The inflectional morphological priming observed was not significantly different from the derivational priming, but as Figure 5.1 above shows, there was a trend towards a greater magnitude of priming in the inflectional condition, in both language groups. Thus, L2 participants appear similar to L1 participants in terms of their ability to make use of morphological structure, both derivational and inflectional, during lexical access.

6.4.2 Research Question 2

Research question 2 asks whether L2 learners' morphological sensitivity is limited to the lexical level or whether L2 learners make use of this morphological information during sentence processing. The results of the self-paced reading task suggest that L2 learners'

knowledge of Arabic morphology is not limited to the lexical level. Rather it appears to be integrated and available to L2 learners during online sentence processing. Analyses of the self-paced reading task data indicated significant effects of all three error conditions (derivational, inflectional and semantic). That is to say, RTs in all three error conditions showed significant slowdowns relative to the baseline condition at positions between the error word and the spillover region (i.e., positions 0 through 2), whereas no condition differed significantly from baseline before the error word (i.e., position -1). The effect of language group points to the fact that L2 learners generally had slower reading times, as is apparent from Figures 5.2 and 5.3 above. There was a significant interaction between language group and condition; however, when L1 and L2 data were analyzed separately, RTs in all three error conditions remained significantly different from baseline for both groups.

An alternate explanation for L2 participants' slowdowns in the error conditions was considered, namely that they might have slowed down because the verbs in those conditions were not familiar to them. This explanation is unlikely because verbs across all four conditions had comparable frequencies and should have been equally likely to be familiar to L2 participants. However, future research should include a vocabulary survey of the verbs used in all conditions in order to rule out unfamiliar words as a cause of slower reading times.

6.4.3 Research Question 3

Research question 3 asks, how does automatic or integrated L2 knowledge of morphology compare to explicit, conscious L2 knowledge of morphology? The results of the acceptability judgment task indicate that while L2 learners of Arabic appear equally

sensitive to inflectional and derivational morphological violations so long as that sensitivity is measured online via reading times, when it comes to offline, metalinguistic measures, L2 learners perform better in the inflectional condition. L2 learners were also comparably accurate in the filler condition involving gender agreement errors (recall that the critical inflectional error condition always involved number agreement errors). When L2 learners' relative accuracy judging number and gender agreement is compared to their relative inaccuracy judging both derivational and semantic anomalies, the emerging pattern suggests that L2 participants are more comfortable making judgments about errors of a syntactic nature than about errors involving lexico-semantic mismatches. If L2 participants could point to a clear grammatical error, they tended to mark the sentence unacceptable; in the absence of such categorical evidence, L2 participants assumed a sentence was acceptable. This seeming acceptance bias also accommodates L2 participants' accuracy in the baseline condition, as an acceptance judgment was always the correct choice here.

6.5 Discussion of Theoretical Approaches

Before discussing the theoretical consequences of the current study, this section will briefly review the three theoretical accounts proposed to explain L2 learners' comparably less native-like behavior surrounding morphology: the Combinatorial Entries Hypothesis, the Uninterpretable Features Hypothesis, and the Sentence-Level Dependencies Hypothesis.

The Combinatorial Entries Hypothesis describes an account first put forth by Silva and Clahsen (2008), who found that L2 learners of English exhibited priming between derived forms in a lexical decision task, when they showed none for inflected forms.

Silva & Clahsen argued that derivational morphology is more likely to exhibit priming in L2 behavioral tasks because derived forms get their own lexical entries in the lexicon, and such lexical entries are addressable with the declarative memory on which L2 learners rely. Silva and Clahsen refer to these entries as "combinatorial entries" because, although they can be retrieved full-form, they subsume the sublexical structure of a derived form. As inflectional morphemes are stored in separate entries from the stems they modify, they cannot be retrieved as easily.

The Combinatiorial Entries Hypothesis predicts that L2 participants should be able to make use of derivational morphology during lexical access but not inflectional morphology, thus it predicted that L2 participants in the current study should have shown greater priming in the derivational than in the inflectional condition of the lexical decision task. As it is specifically a claim about lexical representations, it did not make predictions about the tasks involving sentences: the SPR and the AJT.

The Uninterpretable Features Hypothesis, by contrast, suggests that non-native-like L2 behavior involving inflectional morphemes arises not from the nature of the lexical entries that house them but rather from the features they encode. In a pair of studies discussed in section 2.3 above, Jiang (2004; 2007) found L2 English learners to be insensitive to errors involving plural –s. He explained these results in terms of feature interpretability; the L2 English learners in his studies were L1 speakers of Chinese, a language which only rarely makes use of morphological plural marking. This mismatch combined with the fact that plural marking in English is often redundant (in the sense that the information it encodes tends to be recoverable from other sources) can work together

to make a morpheme like –s "invisible" to L2 learners. This invisibility results in selective integration, whereby certain L2 morphemes remain unacquirable.

The Uninterpretable Features Hypothesis predicts that morphemes that encode unfamiliar and/or redundant features should be difficult or impossible to fully integrate into a second language grammar. Thus, this hypothesis predicted that L2 participants in the current experiment would be more sensitive to Arabic number agreement errors (familiar from English grammar) than to derivational verb pattern errors (unfamiliar from English) during the self-paced reading task. The Uninterpretable Features Hypothesis also predicted that L2 participants would be more accurate about number agreement than gender agreement or derivational verb patterns in the acceptability judgment task because, while the Uninterpretable Features Hypothesis is not a claim about metalinguistic knowledge, it nevertheless stands to reason that L2 learners should be more accurate when judging linguistic phenomena that are part of their automatized L2 competence (as opposed to something they have to monitor consciously).

The third alternative explored, the Sentence-Level Dependencies Hypothesis, suggests that it is not specific unfamiliar features that lead to non-native-like morphological behavior in an L2, but rather sentence-level dependencies in general. This account is related to (but not as strong a claim as) Clahsen and Felser's (2006) Shallow Structures Hypothesis. The SSH maintains that L2 learners lack capacity for rule-based processing in the context of sublexical structures as well as syntactic structures. The alternate possibility explored in the current study is that L2 learners are able to engage in rule-based processing at the sublexical level, but that this ability breaks down at the syntactic level. Clahsen has suggested that the difficulty of processing sentence-level dependencies

may correspond to the complexity of the structural relationship being signaled; for instance, he argues that L2 learners tend to make more accurate judgments about English pronominal case marking than about English subject-verb agreement because "SV agreement dependencies span the entire clause (and thus require comparatively complex structural scaffolding, whereas the objective case is assigned locally within the verb phrase" (Clahsen et al., 2010). The Sentence-Level Dependencies Hypothesis would likewise predict more native-like L2 processing of derivational morphemes than of inflectional morphemes, as sentence-level dependencies are more often signaled by inflectional morphemes. Thus, the Sentence-Level Dependencies Hypothesis predicted that L2 participants in the current study would demonstrate sensitivity to derivational and semantic but not inflectional errors during the self-paced reading task, and would make more accurate judgments about derivational and semantic errors than about inflectional errors during the acceptability judgment task. As the next section explains, however, none of these three theoretical claim's predictions were borne out in the current study's results.

6.5.1 Combinatorial Entries Hypothesis

Regarding the Combinatorial Entries Hypothesis, L2 participants in the current study's lexical decision task showed priming between word pairs that shared an inflectional morphological relationship. This inflectional morphological priming was not significantly different from the derivational priming observed, but as Figure 5.1 above shows, there was a trend towards a greater magnitude of priming in the inflectional condition, in both language groups. Not only was there no interaction between language group and condition in the ANOVAs, but analyses of the L2 data alone revealed significant effects

of inflectional as well as derivational priming compared to baseline. If the Combinatorial Entries Hypothesis were true, L2 participants should not have shown priming in the inflectional condition. The current study's lexical decision results for L2 learners of Arabic conflict with what Silva and Clahsen (2008) observed in L2 learners of English, namely that the latter did not exhibit priming between pairs of words that had an inflectional morphological relationship. Freynik, O'Rourke & Gor (submitted) argued conceptually that combinatorial entries, as described by Clahsen & Silva, could not account for the priming observed between derived forms in L2 learners of Arabic:

> If every derived form has its own combinatorial entry which subsumes its sublexical structure (e.g., stem and affixes), it stands to reason that accessing that entry would prime a learner to access that same stem again, and, crucially, this is what most of the studies in the current review of research on L2 derivational morphological processing were testing: RTs to a stem target, following a prime that was a derived form that included that same stem. Combinatorial entries that come with morphological structure packaged inside them are less helpful in explaining priming that spreads from one derived form to another derived form, when neither form completely subsumes the other. (16)

The current study's results support the claim that, at least in Arabic, L2 learners are sensitive to both derivational and inflectional morphological structure. And if learners are sensitive to both kinds of morphological structure, then it cannot be said that their sensitivity hinges on the nature of the lexical entries that serve only derived forms. In other words, the Combinatorial Entries Hypothesis cannot explain the results of the current study.

6.5.2 Uninterpretable Features Hypothesis

Regarding the Uninterpretable Features Hypothesis, L2 participants in the current study's self-paced reading task exhibited slowdowns in both the inflectional and derivational error conditions. Sensitivity to such features in an online measure implies automatized, integrated (as opposed to offline, declarative) knowledge, and weighs in against the Uninterpretable Features Hypothesis. The fact that L2 learners in the current study demonstrated online sensitivity to number agreement errors contrasts with Jiang (2007)'s finding that Chinese learners of English were not sensitive to English number agreement errors when reading time measures were used.

It is important, however, to point out that Arabic number agreement differs from the English number agreement Jiang was investigating in several important ways. For one, English plural –s is one of the morphemes that DeKeyser (2000) lists as vulnerable to age effects due to its lack of perceptual salience. Perceptual salience is the relative noticeability of a linguistic structure; salient structures are easier to perceive. Numerous researchers have suggested that perceptually salient morphemes tend to be acquired earlier than less salient ones (Brown, 1973; Gass & Selinker, 1994; Pye, 1980; Slobin, 1971). However, Goldschneider and DeKeyser (2001) were the first to rigorously operationalize the construct of perceptual salience in such a way that different morphemes could be quantitatively compared with respect to it. Goldschneider and DeKeyser identified five subcomponents of perceptual salience: phonetic substance, syllabicity, relative sonority, stress, and serial position.

Phonetic substance refers to the number of phones that make up a morpheme (for forms with allomorphs, Goldschneider and DeKeyser averaged the number of phones among all

allomorphs). Morphemes with more phones should be more salient than those with fewer. Syllabicity is a binary quality indicating whether or not a given morpheme includes a vowel; morphemes with vowels should be more salient than those without. Relative sonority refers to how phonologically sonorous a given morpheme is. Goldschneider and DeKeyser used Laver's (1994) sonority hierarchy to calculate relative sonority. Laver's sonority hierarchy ranks phones on a scale from 1 to 9, where 1 is least sonorous (stops) and 9 is most sonorous (low vowels). A morpheme's sonority is the sum of sonority rank values of all the phones that comprise that morpheme; more sonorous morphemes should be more salient than less sonorous ones. Stress indicates whether the morpheme in question receives lexical stress; stressed morphemes should be more salient than unstressed ones. Serial position refers to where the morpheme appears with respect to the stem; Goldschneider and DeKeyser did not elaborate on the specifics of serial position's contribution to perceptual salience because all the morphemes they examined were English suffixes (i.e., position final). However, in discussions of other linguistic phenomena DeKeyser has clarified that, with respect to serial position, continuous morphemes should be more salient than discontinuous ones, and among discontinuous morphemes, circumfixes should be more salient than infixes (personal communication, March 23, 2015).

In light of this operationalization of perceptual salience, Arabic plural marking is more salient than English plural marking. All markers of Arabic plurality involve long vowels, giving them higher ranking in terms of phonetic substance, syllabicity and relative sonority than the English plural –s morpheme that Jiang (2007) examined. Further, Jiang's L2 learners of English spoke Chinese as an L1, and Chinese rarely instantiates

morphemic plural marking. Conversely, both English and Arabic exhibit morphemic plural marking, so it is at least possible that English speaking learners of Arabic are able to transfer their expectation of plural marking and number agreement from English to Arabic. Thus, as Arabic plural agreement involves a salient marker of a familiar feature, one might expect it to be comparably well-integrated into the interlanguages of learners whose L1 is English.

That L2 participants should be sensitive to violations of Arabic verbal pattern morphology, which has no morphological equivalent in English, is harder to explain through appeals to transfer. Again, the claim here is about the form-meaning correspondence between the verbal patterns and the event structures they signal; it is this mapping that should be unfamiliar to English speakers and not the event structures themselves. As the apparent learnability of Arabic derivational morphology by L1 speakers of English cannot be explained by transfer, L2 learners' sensitivity to violations in this condition constitutes evidence against the Uninterpretable Features Hypothesis. The same caveat about perceptual salience is, however, also relevant in the context of Arabic verbal patterns. The verbal patterns vary in terms of their degrees of phonetic substance, sonority, stress and serial positions, and it is difficult to say how to rank a more sonorous form that involves only infixing against a less sonorous form that includes circumfixing, just for instance. The current study did not manipulate perceptual salience of the derivational forms in the SPR or AJT tasks (though phonetic substance and serial position were balanced across sestets in the LDT). Future research using the verbal patterns could shed light on the way serial position interacts with other subcomponents of perceptual salience in discontinuous morphemes.

6.5.3 Sentence-Level Dependencies Hypothesis

Regarding the Sentence-Level Dependencies Hypothesis, L2 participants in the acceptability judgment task made accurate judgments about subject-verb agreement errors involving both number (in the critical inflection condition) and gender (in a filler condition). As both number and gender agreement involve sentence-level dependencies (of the kind that Clahsen et al. (2010) cited as difficult for L2 learners to make accurate judgments about, in comparison to less syntactically complex phenomena like case-marking), the Sentence-Level Dependencies Hypothesis cannot account for L2 participants' relative accuracy in these conditions compared to Derivational and Semantic conditions.

Rather, the relevant difference between number and gender agreement on the one side, and derivational and semantic anomalies on the other, is not the linear or even the structural distance the dependency spans (all three conditions involve a mismatch between an adjacent subject-verb pair). Rather the relevant difference between the conditions is the nature of that mismatch. What the semantic and derivational conditions have in common is the lexico-semantic nature of the subject-verb mismatch. The current study found that L2 participants were less willing to judge a sentence unacceptable on lexico-semantic grounds than on strictly grammatical grounds.

One important consideration for acceptability judgment task design is balancing the relative difficulty of the judgments required across conditions, because the easiest error to identify tends to form a kind of reference point against which participants might judge items that are "less wrong" acceptable. Future comparisons of Arabic derivational and inflectional morphology might do better to focus on more subtle inflectional phenomena.

6.6 Theoretical review and conclusions

To review, the current study found that L2 learners of Arabic exhibited comparably native-like behavior regarding Arabic inflectional morphology across three tasks: a primed lexical decision task, a self-paced reading task and an acceptability judgment task. When it came to Arabic derivational morphology, those same learners showed significant priming between related pairs in the lexical decision task, and significant slowdowns in response to violations during the self-paced reading task, but performed at below chance accuracy when asked to make judgments about those same violations in an acceptability judgment task. This contrast between L2 participants' seeming sensitivity to derivational violations during the self-paced reading task versus their inaccuracy during the acceptability judgment task echoed their performance surrounding semantic violations in both tasks. That is, they likewise evinced slowdowns in response to semantic violations during the SPR, but were unwilling to judge semantically anomalous sentences unacceptable during the AJT, suggesting that, while learners are sensitive enough to all three kinds of violations (inflectional, derivational and semantic) to read more slowly when they encounter them, they are nevertheless not certain enough about violations involving lexico-semantic mismatch to judge them unambiguously unacceptable.

Revisiting the theoretical explanations of L2 morphological deficiency laid out in Section 2.2, it is important to recall that all three were proposed to explain a trend whereby L2 learners (mostly of Indo-European languages) appeared to acquire derivational morphology more quickly and accurately than they did inflectional morphology. As such, none of these frameworks can fully account for the results of the current study in which L2 learners of Arabic demonstrated equal or better command of inflectional morphology

compared to derivational morphology across three experimental tasks. It remains possible that one or more of the hypotheses discussed can adequately account for the patterns observed in L2 acquisition of Indo-European morphologies. Semitic derivational morphology differs from Indo-European in a number of significant ways, including its form, its productivity and its distribution (i.e., Arabic root and pattern morphemes are discontinuous, productive and ubiquitous in a way that few if any Indo-European derivational morphemes are).

Of course, part of the motivation for the current study was the opportunity to examine the acquisition of a system that stretches the bounds of what we usually mean when we talk about derivational morphology. It is plausible that learners of a language like German, for instance, might learn German derived forms in a case-by-case way that differs from the way they learn inflected forms because German derivational morphemes are generally neither as productive nor as predictable in terms of their semantic contributions (compared to German inflections). Arabic verbal pattern morphemes, by contrast, are productive in a way that is more typical of the inflectional systems of Indo-European languages. The sheer productivity of this system requires some degree of generalization across cases[4]. However, the semantic contribution of a given verbal pattern morpheme is not always predictable; in this way, Arabic pattern morphemes are like the derivational morphemes of any other language. The conjecture here is that when you take the relative semantic opacity common to systems of derivational morphology and combine it with the extreme productivity found in Semitic verbal patterns, it stands to reason that you wind

[4] This is evident from the fact that, unlike in Indo-European, morphological decomposition in Semitic languages is mandatory even when that morphology is semantically opaque (Boudelaa & Marslen-Wilson, 2000; Bick, Goelman & Frost, 2011).

up with a learnability profile like the one observed here: L2 learners' grasp of both derivational and inflectional systems look similar until the task requires a conscious judgment, whereupon it becomes apparent that learners are more confident about the grammatical appropriateness of inflectional morphemes than about the semantic appropriateness of derivational morphemes. The inflectional system is simply more predictable and tightly constrained. This remains a post-hoc explanation of the pattern observed. Additional research manipulating the relative productivity and semantic transparency of the morphemes in question will be necessary to pinpoint whether these features of Arabic morphology are the relevant ones in explaining its relative learnability (or lack thereof). Research into the L2 acquisition of the morphologies of languages in other typological families can likewise shed light on this question.

7 Future Directions

One problem with the current study was the lack of a vocabulary survey after the acceptability judgment task to ascertain that participants were equally familiar with all the verbs across conditions in both that task and the self-paced reading task. Thus, a clear first step is to replicate these two experiments with a vocabulary survey afterwards. Though evidence has already been discussed suggesting that learners would be unlikely to be significantly more familiar with verbs in the baseline than in the derivational and semantic conditions of the self-paced reading and acceptability judgment tasks, a vocabulary survey would be the only way to know this for certain.

Another direction for future research concerns event-related potentials. As the Arabic derivational system exhibits some features associated with the derivational systems of other languages (gradations of semantic transparency) and some features more often typical of other inflectional systems (extreme productivity, mandatory decomposition during processing) it is a valid candidate for the focus of an ERP study. ERP measures are useful for examining phenomena at the borders between categories of linguistic processing. They index both the timing and degree of neural activation during language processing, and as different ERP components have been described as pertaining to functionally different stages of language processing, they can lend additional insight in cases where phenomena seem to straddle categories. Results of the current study suggested that in some cases the derivational error conditions patterned with the semantic error conditions (L1 participants responded more slowly to both in the self-paced reading task, L2 participants made inaccurate judgments about both in the acceptability judgment task). ERP data could lend an informative layer to the picture of Semitic derivational

morphological processing that is developing, and speak to the question of whether it is qualitatively more like inflectional morphological processing, semantic processing, or neither.

Additional features that are relevant for research into second language acquisition include the fact that ERPs can highlight both quantitative and qualitative differences between L1 and L2 learners' processing (e.g., differences in latencies or amplitudes of similar waveforms may point to quantitative differences in processing whereas different patterns altogether may point to qualitative differences). Such data could shed additional light on the question of why L2 learners might appear sensitive to morphological errors when their acceptability judgments about them are wildly inaccurate.

References

Alegre, M., and Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language,* 40, 41–61.

Baayen, R. H., Wurm, L. H., and Aycock, J. (2007). Lexical dynamics for low-frequency complex words: A regression study across tasks and modalities. *The Mental Lexicon,* 2(3), 419-463.

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

Basnight-Brown, D.M., Chen, H., Hua, S., Kostic, A., and Feldman, L.B. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Does sensitivity to orthographic similarity vary with language experience? *Journal of Memory and Language*, 57, 65-80.

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

Bick, A., Goelman, G., and Frost, R. (2008). Neural correlates of morphological processes in Hebrew. *Journal of Cognitive Neuroscience*, *20*(3), 406-420.

Boudelaa, S., and Marslen-Wilson, W. D. (2000). Non-concatenative morphemes in language processing: Evidence from Modern Standard Arabic. In *ISCA Tutorial and Research Workshop (ITRW) on Spoken Word Access Processes.*

Boudelaa, S. and Marslen-Wilson, W. D. (2001). Morphological units in the Arabic mental lexicon, *Cognition*, 81, 65-92.

Boudelaa, S. and Marslen-Wilson, W. D. (2004). Allomorphic variation in Arabic: implications for lexical processing and representation. *Brain and Language*, 90, 106-116.

Boudelaa, S. and Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, 20, 207-260.

Boudelaa, S. and Marslen-Wilson, W. D. (2011). Productivity and priming: Morphemic decomposition in Arabic. *Language and Cognitive Processes*, 26, 624-652.

Boudelaa, S., and Marslen-Wilson, W. D. (2000). Non-concatenative morphemes in language processing: evidence from Modern Standard Arabic. *Proceedings of the Workshop on Spoken Word Access Processes,* 1, 23-26.

Bozic, M., and Marslen-Wilson, W. (2010). Neurocognitive contexts for morphological complexity: dissociating inflection and derivation. *Language and Linguistics Compass,* 4(11), 1063-1073.

Brown, R. (1973). *A first language.* Cambridge, MA: Harvard University Press

Buckwalter, T., and Parkinson, D. (2014). *A frequency dictionary of Arabic: Core vocabulary for learners*. Routledge.

Clahsen, H. and Ikemoto, Y. (2012). The mental representation of derived words: an experimental study of –sa and –mi nominals in Japanese. *The Mental Lexicon*, 7(2), 148-182.

Clahsen, H. and Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27: 3-42.

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

Clahsen, H., Sonnenstuhl, I., Blevins, J. (2003). Derivational morphology in the German mental lexicon: a dual mechanism account. In H. Baayen and R. Schreuder (Eds.), *Morphological structure in language processing*. Mouton de Gruyter: Berlin, pp. 125-155.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in second language acquisition*, *22*(04), 499-533.

Deutsch, A. (1998). Subject-predicate agreement in Hebrew: Interrelations with semantic processes. *Language and Cognitive Processes*, 13(5), 575-597.

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

Doron, E., and Hovav, M. R. (2009). A unified approach to reflexivization in Semitic and Romance. *Brill's Journal of Afroasiatic Languages and Linguistics*,*1*(1), 75-105.

Drews, E. and Zwitserlood, P. (1995). Morphological and orthographic similarity in visual word recognition. *Journal of Experimental Psychology; Human, Perception and Performance*, 21, 1098-1116.

Eckman, F. (1994). The competence-performance issue in second-language acquisition theory: A debate. In *Research Methodology in Second-Language Acquisition* ed. by E. Tarone, S. Gass and A. Cohen, 3 - 15. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Feldman, L. B., Kostic, A., Basnight-Brown, D. M., Đurdevic, D. F., and Pastizzo, M. J. (2010). Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism: Language and Cognition*, *13*(02), 119-135.

Frauenfelder, U.H., and Schreuder, R. (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G.E. Booij and J. van Marle (Eds), *Yearbook of morphology 1991*, 165–183. Dordrecht: Foris.

Freynik, S., Gor, K and O'Rourke, P. (2015). *L2 processing of Arabic derivational morphology*. Manuscript submitted for publication.

Frost, R., Forster, K.I., and Deutsch, A. (1997). What can we learn from the morphology of Hebrew: a masked priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 829-856.

Gass, S. M., and Selinker, L. (1994). *Second language acquisition: An introductory course.* Hillsdale, NJ: Erlbaum

Glanville, P. J. (2012). *The Arabic verb: Root and stem and their contribution to verb meaning.* (Doctoral dissertation). The University of Texas at Austin, Austin, Texas.

Goldschneider, J. M., and DeKeyser, R. M. (2001). Explaining the "Natural Order of L2 Morpheme Acquisition" in English: A Meta-analysis of Multiple Determinants. *Language Learning*, 51(1), 1-50.

Gor, K. (2010). Beyond the obvious: Do second language learners process inflectional morphology? *Language Learning*, 60.1, 1-20. Introduction to the thematic issue. Guest editor K. Gor.

Gor, K., and Jackson, S. (2013). Morphological decomposition and lexical access in a native and second language: A nesting doll effect. *Language and Cognitive Processes*, *28*(7), 1065-1091.

Gor, K., and Cook, S. (2010). Non-native processing of verbal morphology: In search of regularity. *Language Learning*, 60.1, 88-126.

Holes, C. (1995). *Modern Arabic: structures, functions, and varieties*. London: Longman.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, 22, 369–397.

Jackson, C. (2008). Proficiency level and the interaction of lexical and morphosyntactic information during L2 sentence processing. *Language Learning*, 58(4), 875-909.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25 (4), 603-634.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, 57, 1–33.

Johnson, J. S., and Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60-99.

Kim, S-Y., Wang, M., and Ko, I-Y. (2011). The processing of derivational words in Korean-English bilingual readers. *Bilingualism: Language and Cognition 14* (4), 473-488.

Kirkici, B., and Clahsen, H. (2013). Inflection and derivation in native and non-native language processing: Masked priming experiments on Turkish. *Bilingualism: Language and Cognition*, *16*(04), 776-791.

Kroll, J. F., and Stewart, E. (1994). Category interferences in translation and picture naming: Evidence for asymmetric connection between bilingual memory representation. *Journal of Memory and Language, 33,* 149-174.

Kuperman, V., Bertram, R., and Baayen, R. H. (2010). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language*, *62*(2), 83-97.

Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania working papers in linguistics*, *4*(2), 14.

Marantz, A. (2001). *Words*. Unpublished manuscript, MIT.

Marslen-Wilson, W. (2007). Processes in language comprehension. The Oxford handbook of psycholinguistics, 11, 175.McKinnon, R. (1996). Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials. *Language and Cognitive Processes*, 11(5), 495-524.

Marslen-Wilson, W., Ford, M., Older, L., and Zhou, X. (1996). The combinatorial lexicon: Priming derivational affixes. In G. Cottrell (Ed.), *Proceedings of the 18th Annual Conference of the Cognitive Science Society* (pp. 223–227). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Marslen-Wilson, W., Hare, M., and Older, L. (1993). Inflectional morphology and phonological regularity in the English mental lexicon. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society.* (pp. 693-698). Hillsdale NJ: Lawrence Erlbaum Associates.

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

McLaughlin, J., Osterhout, L., and Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703-704.

Monsell, S. (1985). Repetition and the lexicon. In A. W. Ellis (Ed.), *Progress in the psychology of language* (pp.147-195). London: Erlbaum.

Murphy, V. A. (1997). The effect of modality on a grammaticality judgement task. *Second Language Research*, *13*(1), 34-65.

Neubauer, K. and Clahsen, H. (2009). Decomposition of inflected words in a second language: An experimental study of German participles. *Studies in Second Language Acquisition 31*, 403-35.

Niswander, E., Pollatsek, A., and Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes*, *15*(4-5), 389-420.

Niswander-Klement, E., and Pollatsek, A. (2006). The effects of root frequency, word frequency, and length on the processing of prefixed English words during reading. *Memory and Cognition, 34*(3), 685-702.

O'Rourke, P. L., and Van Petten, C. (2011). Morphological agreement at a distance: Dissociation between early and late components of the event-related brain potential. *Brain Research, 1392*, 62-79.

Orsolini, M., and Marslen-Wilson, W. (1997). Universals in morphological representation: Evidence from Italian. *Language and Cognitive Processes, 12*(1), 1–47.

Pearlmutter, N. J., Garnsey, S. M., and Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language, 41*(3), 427-456.

Perea, M., Abu Mallouh, R., and Carreiras, M. (2010). The search of an input coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin and Review*, 17, 375-380.

Perea, M., and Lupker, S. J. (2003). Transposed-letter confusability effects in masked form priming. In S. Kinoshita and S. J. Lupker (Eds.), *Masked priming: State of the art* (pp. 97-120). Hove, UK: Psychology Press.

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York: Basic Books.

Portin, M., Lehtonen, M., and Laine, M. (2007). Processing of inflected nouns in late bilinguals. *Applied Psycholinguistics, 28* (1), 135–156.

Portin, M., Lehtonen, M., Harrer, G., Wande, E., Niemi, J., and Laine, M. (2008). L1 effects on the processing of inflected nouns in L2. *Acta Psychologica,* 128 (3), 452–465.

Prunet, J., Beland, R. and Idrissi, A. (2000). The Mental Representation of Semitic Words. *Linguistic Inquiry,* 31, 609-648.

Pye, C. (1980). The acquisition of person markers in Quiche Mayan. *Papers and Reports on Child Language Development*, 19, 53–59.

Randall, B., and Marslen-Wilson, W. D. (1998). The relationship between lexical and syntactic processing. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society,* Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Rastle, K. and Davis, M.H. (2008). Morphological decomposition is based on the analysis of orthography. *Language and Cognitive Processes*, 23(7/8), 942-971

Reid, A., and Marslen-Wilson, W.D. (2000). Organising principles in lexical representation: evidence from Polish. In L.R. Gleitman and A.K. Joshi (Eds.), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*. (pp. 405–410). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Sagarra, N., and Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua, 120*(8), 2022-2039.

Sato, M. (2007). *Sensitivity to syntactic and semantic information in second language sentence processing.* (Unpublished doctoral dissertation). University of Essex.

Schreuder, R. and Baayen, R. H. (1995). Modeling morphological processing. In Feldman, L. B., editor, *Morphological aspects of language processing* (pp. 131–154). Lawrence Erlbaum, Hillsdale, New Jersey.

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

Slobin, D. I. (1971). Data for the symposium. In D. I. Slobin (Ed.), *The ontogenesis of grammar* (pp. 3–14). New York: Academic Press.

Smolka, E., Zwitserlood, P., and Rösler, F. (2007). Stem access in regular and irregular inflection: Evidence from German participles. *Journal of Memory and Language, 57*(3), 325-347.

Sonnenstuhl, I., Eisenbeiss, S., and Clahsen, H. (1999). Morphological priming in the German mental lexicon. *Cognition, 72*, 203–236.

Soveri, A., Lehtonen, M., and Laine, M. (2007). Word frequency and morphological processing in Finnish revisited. *Mental Lexicon, 3*(2), 359–385.

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

Tarone, E. (1988). *Variation in interlanguage.* London: Edward Arnold.

Vincenzi, M. D., Job, R., Di Matteo, R., Angrilli, A., Penolazzi, B., Ciccarelli, L., and Vespignani, F. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language, 85*(2), 280-296.

Whong-Barr, M., and Schwartz, B. D. (2002). Morphological and syntactic transfer in child L2 acquisition of the English dative alternation. *Studies in Second Language Acquisition, 24*(04), 579-616.

Wunderlich, D. (1996). Minimalist morphology: the role of paradigms. In *Yearbook of Morphology 1995* (pp. 93-114). Springer Netherlands.

Appendix A: LANGUAGE BACKGROUND QUESTIONNAIRE

Test ID: _____ **Date:** _____

1. I am right-handed _____ left-handed _____ ambidextrous _____ (**Check one**)
2. I am MALE FEMALE (**Please, circle one of the options)**
3. I am _____ years old
4. I am a FRESHMAN SOPHOMORE JUNIOR SENIOR GRADUATE STUDENT NOT A STUDENT
5. My major is (was) _____
6. My native language is (circle) ARABIC ENGLISH OTHER (Specify) _____
7. The second language (L2) I spoke/learned was ARABIC ENGLISH OTHER (Specify) _____

8	I started learning ENGLISH when I was_____y.o.	I started learning ARABIC when I was _____y.o.
9	I started learning ENGLISH: □ At home □ In school □ In college/university □ In the community	I started learning ARABIC: □ At home □ In school □ In college/university □ In the community
10	I started learning ENGLISH: □ In an ENGLISH-speaking country In a ARABIC-speaking country	I started learning ARABIC: □ In an ENGLISH-speaking country □ In a ARABIC-speaking country
11	I had formal instruction in ENGLISH in grade school for _____mnths/yrs in college for_____mnths/yrs other (specify)_____mnths/yrs □	I had formal instruction in ARABIC in grade school for _____mnths/yrs in college for_____mnths/yrs other (specify)_____mnths/yrs
12	I lived in an ENGLISH-speaking country for _____ (mnths/ yrs) when I was _____y.o., for(mnths/ yrs) _____when I was ____y.o., etc. (list all your visits) With a total of _____mnths/yrs	I lived in a ARABIC-speaking country for _____ (mnths/ yrs) when I was _____y.o., for(mnths/ yrs) _____when I was ____y.o., etc. (list all your visits) With a total of _____mnths/yrs

13. List what percentage of the time you have been exposed to each language:

…WHEN I WAS **0-5** YEARS OLD	***ARABIC***	0 - 20%	21 - 40%	41 - 60%	61 - 80%	81 - 100%
	ENGLISH	81 - 100%	61 - 80%	41 - 60%	21 - 40%	0 - 20%
	Put a checkmark here →					
…WHEN I WAS **6-10** YEARS OLD	***ARABIC***	0 - 20%	21 - 40%	41 - 60%	61 - 80%	81 - 100%
	ENGLISH	81 - 100%	61 - 80%	41 - 60%	21 - 40%	0 - 20%
	Put a checkmark here →					
…WHEN I WAS **11-15** YEARS OLD	***ARABIC***	0 - 20%	21 - 40%	41 - 60%	61 - 80%	81 - 100%
	ENGLISH	81 - 100%	61 - 80%	41 - 60%	21 - 40%	0 - 20%
	Put a checkmark here →					
…WHEN I WAS **16 -20** YEARS OLD	***ARABIC***	0 - 20%	21 - 40%	41 - 60%	61 - 80%	81 - 100%
	ENGLISH	81 - 100%	61 - 80%	41 - 60%	21 - 40%	0 - 20%
	Put a checkmark here →					
…WHEN I WAS **21** AND OLDER	***ARABIC***	0 - 20%	21 - 40%	41 - 60%	61 - 80%	81 - 100%
	ENGLISH	81 - 100%	61 - 80%	41 - 60%	21 - 40%	0 - 20%
	Put a checkmark here →					

14. Using the following scale, rate your language proficiency in 1) ENGLISH, if you are a native speaker of ARABIC, 2) ARABIC, if you are a native speaker of ENGLISH, 3) BOTH if you are a heritage speaker

ENGLISH Minimal ----------Native-like

Speaking	1	2	3	4	5	6	7	8	9	10
Pronunciation	1	2	3	4	5	6	7	8	9	10
Listening	1	2	3	4	5	6	7	8	9	10
Reading	1	2	3	4	5	6	7	8	9	10
Writing	1	2	3	4	5	6	7	8	9	10
Grammar	1	2	3	4	5	6	7	8	9	10

ARABIC Minimal ----------Native-like

Speaking	1	2	3	4	5	6	7	8	9	10
Pronunciation	1	2	3	4	5	6	7	8	9	10
Listening	1	2	3	4	5	6	7	8	9	10
Reading	1	2	3	4	5	6	7	8	9	10
Writing	1	2	3	4	5	6	7	8	9	10
Grammar	1	2	3	4	5	6	7	8	9	10

Appendix B Consent Form

Project Title	*Understanding Arabic Words Alone and in Context*
Purpose of the Study	*This research is being conducted by **Dr. Kira Gor** and **Suzanne Freynik** at the University of Maryland, College Park. We are inviting you to participate in this research project because you are an adult native speaker of English who has studied Arabic as a second language, or because you are an adult native speaker of Arabic. The purpose of this research project is to determine how native English speakers who learn Arabic as adults compare to native speakers of Arabic in the ways they understand Arabic words alone and in context.*
Procedures	*The procedures involve (a) looking at strings of Arabic letters on a laptop screen and then pushing a button to indicate whether a given string is a real Arabic word, (b) reading Arabic sentences on a laptop screen and answering Yes/No questions about some of them, and (c) reading Arabic sentences on a laptop screen and making judgments about how grammatical and/or sensical they are. You may also be asked to answer fill-in-the-blank questions in Arabic, and to translate a list of Arabic words, as well as to fill out a questionnaire about your language learning experiences (e.g., At what age did you begin learning Arabic?, How many years of formal Arabic instruction have you had?). The experiment will involve two separate sessions, and take no longer than 3 hours total.*
Potential Risks and Discomforts	*The risks of these research methods are minimal, but include the following: boredom or sleepiness, and risks normally associated with using a computer monitor and keyboard, such as eye-strain. The tasks are self-paced and you will have the opportunity to take breaks in order to mitigate these risks.*
Potential Benefits	*This research is not designed to help you personally, but the results may help the investigator learn more about how adults' language learning compares with children's language learning. We hope that, in the future, other people might benefit from this study through improved understanding of language acquisition.*
Confidentiality	*Any potential loss of confidentiality will be minimized by through the following means: data collected for each participant will be assigned a number and will subsequently be identified only by that number. All the data will be stored on password-protected computer files, and attendant documents such as consent forms will be stored in a locked cabinet in a locked office. Only approved researchers (Suzanne Freynik, Dr. Kira Gor and Dr. Polly O'Rourke) will have access to this data.* *If we write a report or an article about this research project, your identity will be protected to the maximum extent possible; your name and/or initials will never be used, and any description of personal information will be limited to including your age in a description of the average participant age, and your gender in a description of the distribution of participant gender. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you*

	or someone else is in danger or if we are required to do so by law.
Compensation	*You will receive $10 for the first session and $30 for the second session of this study. You will be responsible for any taxes assessed on the compensation (see next page).* ☐ *Check here if you expect to earn $600 or more as a research participant in UMCP studies in this calendar year. You must provide your name, address and SSN to receive compensation.* ☐ *Check here if you do not expect to earn $600 or more as a research participant in UMCP studies in this calendar year. Your name, address, and SSN will not be collected to receive compensation.*
Right to Withdraw and Questions	*Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify.* *If you decide to stop taking part in the study, if you have questions, concerns, or complaints, or if you need to report an injury related to the research, please contact one of the investigators:* ***Suzanne Freynik*** ***Center for Advanced Study of Language*** ***7005 52nd Avenue*** ***College Park, MD 20742*** ***(e-mail) freynik@umd.edu*** ***(telephone) 570-772-7479*** ***Kira Gor*** ***2106E*** ***Jimenez Hall*** ***University of Maryland*** ***College Park, MD 20749*** ***(email) kiragor@umd.edu*** ***(telephone) 301-405-0185***
Participant Rights	*If you have questions about your rights as a research participant or wish to report a research-related injury, please contact:* **University of Maryland College Park** **Institutional Review Board Office** **1204 Marie Mount Hall**

	College Park, Maryland, 20742 **E-mail: irb@umd.edu** **Telephone: 301-405-0678** *This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects.*	
Statement of Consent	*Your signature indicates that you are at least 18 years of age; you have read this consent form or have had it read to you; your questions have been answered to your satisfaction and you voluntarily agree to participate in this research study. You will receive a copy of this signed consent form.* *If you agree to participate, please sign your name below.*	
Signature and Date	**NAME OF SUBJECT** **[Please Print]**	
	SIGNATURE OF SUBJECT	
	DATE	

Appendix C Debriefing

DEBRIEFING
Understanding Arabic Words Alone and in Context

Thank you for your participation in our research study, *Understanding Arabic Words Alone and in Context.*

I would like to discuss with you in more detail the study you just participated in and to explain exactly what we were trying to study.
Before I tell you about all the goals of this study, however, I want to explain why it is necessary in some kinds of studies to not tell people all about the purpose of the study before they begin.

As you may know, scientific methods sometimes require that participants in research studies not be given complete information about the research until after the study is completed. Although we cannot always tell you everything before you begin your participation, we do want to tell you everything when the study is completed.

We don't always tell people everything at the beginning of a study because we do not want to influence your responses. If we tell people what the purpose of the study is and what we predict about how they will react, then their reactions would not be a good indication of how they would react in everyday situations.

The purpose of this study is to examine how native English speakers who learn Arabic as adults compare to native speakers of Arabic in the ways they understand Arabic words. Arabic words, like some English words, are made up of smaller units called morphemes, but in Arabic they fit together in a way that is very different from the morphemes of most other languages (English morphemes fit one after another like boxcars, whereas many Arabic morphemes interleave together, like teeth in a zipper). Some linguists call these Arabic morphemes "roots" and "patterns".

This study uses three methods called lexical priming, self-paced reading, and acceptability judgment to investigate how, and to what extent, speakers are able to recognize and understand these units that make up Arabic words. In the lexical decision task, you heard a spoken word before you saw each written word on the computer screen. In the trials we are interested in, the spoken word had the same Arabic root as the written word. We wanted to know if hearing a word with the same root would help you recognize the written word faster. In some of the other trials, the spoken word had a similar sound to the written word, or it had a similar meaning. We included these trials so that we could compare what happens when the words share roots with what happens when the words share only sounds or only meaning.

In the self-paced reading task, you read Arabic sentences one word at a time. In the trials we are interested in, the verbs were inappropriate because they had the wrong verbal patterns. In some of the other trials, the verbs had the wrong inflections (e.g., the verb was plural when the

subject was singular), or the whole verb was just an implausible fit for the rest of the sentence. All of these changes might make the sentences harder to understand, so people read the problematic sentences more slowly. We wanted to know if inappropriate verb patterns cause more or less disruption to reading than inappropriate verb inflections, or than verbs that just don't fit at all.

Scientists believe that many important aspects of word processing happen unconsciously. We did not tell you about the specific parts of the words that we were interested in because we wanted to observe these unconscious effects, and we did not want you to consciously look for the roots and patterns. However, we were also interested in how you might react to the sentences when you had more time, and were asked to look for errors. That is, we wanted to compare unconscious knowledge with conscious knowledge. For this reason, we included the acceptability judgment task in the second session, and asked you to make conscious judgments about sentences that were similar to the ones you read during the first session.

If other participants knew the specific purpose of the study, it might affect how they behave, so we are asking you not to share the information we just discussed until after the study is over.

Now that the study has been explained, do you agree to allow the investigator to use the data that we collected from your participation in this study?

I hope you enjoyed your experience. If you have any questions later please feel free to contact us.

Suzanne Freynik
Center for Advanced Study of Language
7005 52nd Avenue
College Park, MD 20742
(e-mail) freynik@umd.edu
(telephone) 570-772-7479

Or

Kira Gor
2106E Jimenez Hall
University of Maryland
College Park, MD 20749
(email) kiragor@umd.edu
(telephone) 301-405-0185

Do you have any other questions or comments about anything you did today or anything we've talked about?

Thank you again for your participation.

Appendix D Lexical Decision Task Master List

#	Target	Deriv	Infl	Phon	Semantic	Base
1	سمع	إستمع	نسمع	أسمار	إذن	قدّر
	to hear	to listen	we listen	brown, olive complected	ear	to estimate
2	بارك	تبارك	يبارك	ترك	صلاة	استوعب
	to bless (fi) sb (of God)	to be praised (of God)	he blesses	leave	prayer	to assimilate, absorb
3	تبع	تابع	تبعوا	تعب	تعقب	همس
	to follow, pursue sth/sb	to follow, monitor sth/sb	they followed	tire, get tired	chase	to whisper
4	حبّ	أحبّ	يحبّ	محاسب	عشق	وزع
	to love, like sb; to want, like sth	to love, like sth/sb; to want sth	he loves	examination, accounting	passion, love	to distribute
5	حدّ	حدّد	حدّوا	حدث	توقف	نام
	to limit (min) sth; to halt, stop (min) sth	to specify sth; to define sth	they limited	happened, occurred	to stop	to sleep
6	حسّ	أحسّ	يحسّ	أحسن	لمس	نظّم
	to feel, sense (bi) sth	to feel, sense (bi) sth or (anna) that	he feels	best	touch, sense	to organize
7	حفظ	إحتفظ	يحفظ	إحتفل	تذكر	انتقم
	to preserve; to memorize	to preserve, keep (bi) sth	he preserves	celebrate	remember	to take revenge
8	تحقّق	إستحقّ	يتحقّق	احتقان	ثبت	تمرّد
	to become reality; to verify (fi) sth	to deserve, merit sth	he verifies	congestion, political tension	prove, fix, confirm	to rebel
9	حكم	تحكّم	يحكم	إحتكار	سيطر	تنوّع
	to govern; to sentence (3ala) sb	to control (fi/bi) sth	he governs	monopoly, hoarding	control	to be of various kinds
10	حوى	إحتوى	نحوي	محور	طوق	وصف

	to contain, include (3ala) sth	to contain, include 3ala sth	we include	axis	to surround, embrace, include	describe
11	نشر	أنشر	ينشر	نشاط	أعلن	أوفد
	to publish	to be spread, to be published	he publishes	activity	advertise, declare	to send a delegation
12	خبّر	أخبر	نخبّر	إختار	إعلام	أنقذ
	to tell sb sth or 3an about sth	to notify, tell sb 3an/ bi about sth	we tell about	choice	to inform	to rescue, save
13	ختم	اختم	نختم	مخاطب	طَبَع	سبح
	to conclude sth; to seal, stamp sth	to finalize (an activity)	we conclude	speaking, conversation	to stamp, print	to swim
14	خرق	إخترق	نخرق	إخترع	خالف	تناول
	to violate (law)	to break into, to traverse	we break the law	invention	to transgress	to take (a meal)
15	دفع	دافع	دفعوا	دفتر	صرف	إهتزّ
	to push; to pay; to compel	to defend 3an sth/sb	they pushed	folder, notebook	to pay	to shake, tremble
16	رجع	تراجع	يرجع	مرجو	عاد	تهذّب
	to return ila	to retreat; to decrease	he returns	requested, wished for	to return	to become refined, educated
17	رعى	راعى	رعينا	رعب	أيد	هاجر
	to protect sb; to sponsor sth/sb	to heed, observe sth; to respect sth	we protected	fright, panic	to support, help	to emmigrate
18	زال	أزال	يزول	زميل	إختفى	نقد
	to disappear, vanish	to disappear, vanish	he disappears	colleague	to disappear	criticize
19	زاد	تزايد	يزيد	تزاوج	مجموع	نسي
	to increase; to exceed	to increase, grow in number	he exceeds	to marry	sum, total	to forget
20	سبّب	تسبّب	نسبّب	مسابق	أحدث	مشى

	to cause, produce, provoke sth	to cause, result in fi / bi sth	we provoke	contest, competition	to provoke, induce	to walk
21	سدّ	سدّد	سدّنا	سواد	قفل	وعى
	to close; to turn off; to pay; to fill	to obstruct; to pay off; to aim, shoot	we turned	darkness, blackness	to close, shut	to be aware
22	سارع	أسرع	يسارع	مسرح	عجل	كتب
	to hurry, hasten ila to a place	to hurry, hasten fi in doing sth	he hurries	theater, stage	to hurry	to write
23	شجع	أشجع	نشجع	شعر	جريء	إقتنع
	to be brave	to encourage	we are brave	to feel	bold, daring	to be convinced
24	سقط	تساقط	يسقط	سقف	وقع	شرب
	to fall; drop, decline	to collapse	he falls	roof, ceiling	to fall	drink
25	أسلم	تسلّم	يأسلم	تسلل	رخا	قرّا
	to surrender, hand over sth	to receive sth; to take on sth	he surrenders	infiltrate	to relinquish	to read
26	سمّى	أسمى	يسمّي	تسامح	دعا	غسل
	to name, designate, call	to name, designate, call	he names	tolerance	call, name	to wash
27	ساهم	اسهم	يساهم	سهر	تعاون	زار
	to participate in, contribute to	to participate, contribute, share	he participates	vigil	to cooperate, collaborate, participate	visit
28	شابه	أشبه	نشابه	اشتباك	ماثل	نادى
	to resemble, be similar to sth/sb	to resemble, look like sth/sb	we resemble	skirmish, clash	to resemble	to call out
29	أشرف	تشرّف	يأشرف	تشريع	أدار	فكّر
	to supervise, manage 3ala sth/sb	to be honored (3ala to meet sb)	he supervises	legislation	to manage	ponder
30	شارك	إشترك	نشارك	إشترى	تعاون	ضحك

	to participate (with sb) fi in	to participate (with sb) fi in	we participate	to buy	to cooperate	laugh
31	انهى	إنتهى	انهوا	نهار	كمل	سافر
	to complete, finish	to end, conclude, finish	they completed	daytime	to complete, supplement	travel
32	صعد	تصعد	يصعد	أصعب	نهض	لمّح
	to rise, go up; to increase	to climb, increase	he goes up	more difficult	to rise, get up	to hint
33	طال	طوّل	طالوا	طفل	وسع	انتصف
	to be lengthy, to take awhile	to prolong, to take time	they took awhile	child	to be wide, to extend	to be in the middle
34	ضمن	ضمّن	ضمنوا	ضمير	شمل	غادر
	to guarantee, insure	to guarantee, insure	they guaranteed	conscience	to include, cover	to leave
35	أضاف	إستضاف	يأضاف	ضعيف	جمع	غنى
	to add sth	to host, invite sb	he adds	weakness	to gather, combine, add	to sing
36	طرق	تطرّق	نطرق	مطروح	ضرب	فضّل
	to knock on (door)	to broach, discuss ila (topic, issue)	we knock	proposed, offered	to hit	to prefer
37	طلع	إطّلع	يطلع	أطلّ	برز	ظنّ
	to appear, emerge; to go out	to examine, peruse 3ala sth	he appears	to overlook, provide view	to emerge, protrude	to think
38	عدا	تعدّى	نعدو	معدني	جرى	افتقر
	to run, race	to go beyond; to infrine 3ala on	we race	metal	to run	to lack
39	عقد	اعتقد	يعقد	معاقب	اجتمع	طوّر
	to hold, convene (meeting)	to believe fi in sth, or bi'inna/anna that	he convenes	punishment, sanction	to meet, have a meeting	to develop
40	طلب	طالب	طلبوا	طبغ	حاجة	فقد

	to request, demand something	to request (bi) something (from somebody)	they requested	to cook	need	to lose
41	علق	تعلّق	يعلق	معالج	ربط	أرسل
	to be pending; to be attached	to be connected bi with sth/sb	he is connected to	treatment	to link, connect	to send, mail
42	عمد	إعتمد	يعمد	تعميم	قصد	آكل
	to do sth deliberately	to depend, rely 3ala on sth/sb	he acts deliberately	publicizing	to mean, aim	to eat
43	تعوّد	إعتاد	نتعوّد	أعوام	إستقرّ	لعب
	to get accustomed 3ala to sth	to get used 3ala to sth	we get accustomed	years	to settle, stabilize	to play
44	عيّن	تعيّن	يعيّن	أعياد	خصّص	طرّ
	to appoint sb; to define sth	to be incumbent 3ala on sb	he appoints	feasts	to specify, assign	to occur
45	خلق	اختلق	نخلق	خليج	صنع	نظر
	to create	to feign, fabricate, invent	we create	gulf	manufacture	to watch
46	غفر	استغفر	نغفر	استغرب	سامح	استند
	to pardon, forgive li sb sth	to beg (God) for forgiveness	we forgive	be surprised	to pardon, permit	to have a basis in
47	غلق	أغلق	يغلق	أغلبية	قفل	طار
	to bolt shut, to close (door)	to lock or bolt shut, to close (door)	he closes	majority	to close, lock	to fly
48	فتح	إفتتح	يفتح	أفراح	باب	استقلّ
	to open sth; to turn on (lights, TV)	to open, inaugurate sth	he opens	joys, celebrations	door	to become independent
49	قبل	استقبل	يقبل	مقبرة	أقر	زيّن
	to accept, recieve; approve	to meet, welcome, greet sb	he accepts	graveyard	to agree, accept	to decorate
50	قتل	قاتل	قتلوا	قليل	مات	عمّم

	to kill sb	(3) to fight sb	they killed	little	to die	to generalize
51	قدم	تقدّم	نقدم	قدوة	جاء	أصرّ
	to arrive, come ila to	to present sth, to advance	we arrive	example, pattern	to come over	to insist
52	قرب	قارب	قربنا	قرن	اتّصل	دهش
	to approach	to come close to sth	we approached	century	to contact, reach, approach	to be amazed
53	قام	قاوم	قامنا	قمر	إرتفع	نظف
	to stand; to carry out b (task)	to resist, oppose sth	we stood	moon	to rise, climb	to clean
54	كثر	أكثر	نكثر	تكثيف	ضاعف	ابتسم
	to be plentiful	to do min sth frequently	we are many	intensifying, compression	to multiply, double	to smile
55	كسب	إكتسب	يكسب	مكتبة	نجح	دوّن
	to gain, achieve, earn sth	to earn, gain, win sth	he earned	library	succeed, achieve	to write down
56	كفل	كفّل	كفلوا	كفاح	رهن	استراح
	to guarantee sth; to support sb	to support, maintain, provide for s b	they guaranteed	struggle	to guarantee	to relax
57	ولى	توالى	يلي	ويل	تالي	جهل
	to follow, come after	to follow in succession	he follows	woe, distress	next	to ignore
58	لحق	ألتحق	يلحق	ملاحظ	إنضمّ	دخّن
	to follow, be attached to	to append bi sth	he follows	noticing	to join, be annexed to	to smoke
59	كلّم	تكلّم	نكلّم	كلّية	قال	هان
	to speak with, talk to	to speak (ma3) with	we speak with	college	to say	to betray
60	لفت	إلتفت	نلفت	ملفّ	جذب	خمن
	to turn sb's attention ila to	to turn around	we direct attention to	file, dossier	to attract, engage	to guess
61	لقى	التقى	يلقي	إلتقاط	صادف	إختفل
	to find; meet,	to meet, encounter bi/ma3 sb	he finds	receiving, taking	to chance, encounter, meet	to disagree

	encounter sb/sth					
62	لاح	لوّح	لاحوا	لحم	ظهر	جنّن
	(u) to appear, loom	to wave ila at sb; to hint at bi sth	they appeared	meat	to appear	to drive crazy
63	مدّ	إمتدّ	نمدّ	مدينة	بسط	ترجم
	to extend sth; to stretch out sth	to extend, reach, spread ila to	we stretch	city	to extend, spread, stretch	to translate
64	مرّ	إستمرّ	يمرّ	تمرين	عبر	رتّب
	(u) to go past; to stop by 3ala (sb's place)	to last; to continue fi doing	he goes past	drill, exercise	to express, cross	to arrange, prepare
65	مسك	أمسك	يمسك	سكين	انتهز	حلم
	to grab, hold sth or bi sth	to hold sth; to refrain from sth	he grabs	knife	to seize	to dream
66	غيّر	تغيّر	يغيّر	تغيب	تبدل	إشتهر
	to change something	to be modified	he changes	to be absent from	to transform, change	to be famous
67	ملك	إمتلك	نملك	تمويل	ساد	اختصر
	to own, possess sth; to control sth	to possess, own sth	we own	financing, funding	to dominate, rule	to abridge, abbreviate
68	ناسب	تناسب	يناسب	منسق	لائم	روى
	to be suitable for sb	to be compatible wa/ma3 with	he is suitable	coordinator	to suit, fit, accommodate	to narrate
69	اثّر	تاثّر	ناثّر	ثورة	نتيجة	استحمّ
	to affect something	to be affected by	we affect	revolution	result, consequence	to bathe
70	وجد	تواجد	نجد	توجّب	إكتشف	حرّر
	to find sth/sb (present tense: there	to be located; to be present	we find	to be necessary	to find, discover	to liberate

	is/ there are)					
71	وفّر	توفّر	نوفّر	موظف	تزويد	دلّ
	to be met, fulfilled fi in sth/sb	to be abundantly available	we provide	employee	supply	to indicate, point to
72	وافق	إتّفق	نوافق	وفاة	تطبق	رسم
	to agree with sb	to agree (ma3) with sb	we agree	death	to match, correspond	to draw
73	عكس	إنعكس	يعكس	كسر	ضد	ناقش
	to reflect, reverse	to be reflected, to have an effect	he reverses	to break	opposed, against, opposite	to discuss
74	مثّل	تمثّل	يمثّل	مثلج	رمز	إنقطع
	to represent, to act for	to be represented, to be seen in	he represents	frozen	symbol,	to be cut off
75	بقي	أبقى	يبقي	بيّن	ظل	سكت
	to remain, continue	to keep something in a state	he continues	to clarify	to remain, stay	to become silent

Appendix E Self-Paced Reading Task Master List

#	Sentence	Base	Deriv	Infl	Semantic
1	الأزهار كانت متفتحة والرائحة ذكّرت المسافر عن رحلته في مصر.	ذكّرت	تذكّرت	ذكّروا	وثقت
	The orange blossoms were open and the smell reminded the traveler of his visit to Egypt.	reminded	remembered	reminded (pl)	trusted
2	إلى الشرق من الحدود، الطريق خرج عن المسار الذي اتبعه من قبل.	خرج	أخرج	خرجوا	اتهم
	To the east of the border, the road deviated from the path that it followed before.	deviated	expelled	deviated (pl)	accused
3	الأسبوع الماضي، القصّة صدرت في مجموعة كتابة مماثلة.	صدرت	أصدرت	صدروا	غضبت
	Last week, the story was published in a collection of similar writing.	was published	published	were published	got angry
4	التحفيظ كان صعب، ولكن التكرار عزّز الذاكرة بشكل فعال وكاف.	عزّز	عزّ	عزّزوا	فاح
	Memorization was difficult, but repetition strengthened memory in an effective enough way.	strengthened	was strong	strengthened (pl)	wafted
5	بعد الفيضان العسير، الطائرة أحضرت الأكل للمساكين في الجزيرة.	أحضرت	حضرت	أحضروا	شخصت
	After the serious flood, the plane brought food to the poor people on the island.	brought	attended	brought (pl)	diagnosed
6	بعد درس اليوم، الأستاذ وضّح الواجب وكتبه على السبورة.	وضّح	أتضح	وضّحوا	انصهر
	After today's lesson, the teacher clarified the homework and wrote it on the board	clarified	became clear	clarified (pl)	melted
7	بعيدا عن البلدة، القناة نقلت الماء من النهر إلى المعمل.	نقلت	انتقلت	نقلوا	دقت
	Far from the town, a canal transferred water from the river to the factory.	transferred	moved over	transferred (pl)	knocked
8	بقية الطعام كانت سيئة، ولكن الشوربة أشبعت جوع الضيوف في المنزل.	أشبعت	شبعت	أشبعوا	نطقت
	The rest of the food was bad, but the soup satisfied the hunger of the guests in the house.	satisfied	was satisfied	satisfied (pl)	pronounced
9	خارج البيت، الحديقة ارتكزت على شجرة التفاح القديمة.	ارتكزت	ركّزت	ارتكزوا	صرخت
	Outside the house, the garden centered around the old apple tree.	centered around	focused	centered around (pl)	shouted
10	خلال القرن الماضي، النهر حمل البضائع إلى البلدة التالية في سفن.	حمل	أحمل	حملوا	أراد
	During the last century, the river carried goods to the next city on boats.	carried	loaded	carried (pl)	wanted
11	خلال المناقشة أمس، الأمر تقرّر بدون عنف أو كلمات قاسية.	تقرّر	قرّر	تقرّروا	بكى

	During the meeting yesterday, the matter was resolved without violence or harsh words.	was resolved	decided	were resolved	cried
12	شكل الكرسي كان غريبا، ولكن اللون انطبق مع بقية الأثاث في الغرفة.	انطبق	طبّق	انطبقوا	شكا
	The shape of chair was strange, but the color conformed with the rest of the furniture in the room.	conformed	implemented	conformed (pl)	complained
13	طوال فترة الأعياد، المدير قلّل الساعات التي يكون الدكان مفتوح فيها.	قلّل	قلّ	قلّلوا	ذاب
	During the holidays, the manager reduced the hours when the store was open.	reduced	shrunk (himself)	reduced (pl)	dissolved
14	عندما أصبح الخبز غالي، السعر أثار احتجاج في الحي الفقير.	أثار	ثار	أثاروا	لمع
	When bread became expensive, the price stirred up a demonstration in the poor neighborhood.	stirred up	revolted	stirred up (pl)	shined
15	في الحفلة ليلة أمس، الموسيقى أسعدت النساء رغم إنهنّ تعبانات.	أسعدت	سعدت	أسعدوا	لامت
	At the party last night, the music made the ladies happy even though they were tired.	pleased	became happy	pleased (pl)	blamed
16	في الطريق وسط الجبال، الحادث أوقف الازدحام لوقت طويل.	أوقف	وقف	أوقفوا	ارتاح
	On the road through the mountains, the accident detained the traffic for a long time.	detained	stopped (itself)	detained (pl)	relaxed
17	في بداية الربيع، الرطوبة لزمت الأرز الذي يزرع هنا.	لزمت	التزمت	لزموا	أملت
	In the beginning of the spring, the humidity was necessary for the rice that grows here.	was necessary	committed to	were necessary	hoped
18	في شمال البلاد، البرد حذّر الناس من الشتاء القادم.	حذّر	حاذر	حذّروا	حسد
	In the north of the country, the cold warned the people of the coming winter.	warned	was careful	warned (pl)	envied
19	في صحيفة اليوم، المقالة عرّفت الرئيس الجديد بوصف مناسب.	عرّفت	تعرفت	عرّفوا	تدفقت
	In today’s newspaper, the article introduced the new president with an appropriate description.	introduced	met	introduced (pl)	dripped
20	في نهاية هذا الفصل، الواجب شغل الطلاب حتى أنهم سهروا كل الليلة.	شغل	اشتغل	شغلوا	تحمس
	At the end of the semester, the homework preoccupied the students until they stayed up all night.	preoccupied	was worried	preoccupied (pl)	got excited
21	كان عادي ورخيص الثمن، ولكن السكين خدم الجزار في شغله بدون مشاكل.	خدم	استخدم	خدموا	تعهد
	It was ordinary and cheap, but the knife served the butcher in his work without problems.	served	used	served (pl)	pledged

22	كان هناك تسرب في السطح، و الماء ملأ الدلو الذي كان تحته.	ملأ	امتلأ	ملأوا	عانى
	There was a leak in the roof, and the water filled the bucket that was under it.	filled	was filled up	filled (pl)	suffered
23	كل مناطق البلاد جميلة، ولكن الصحراء تميّزت بجمالها وطبيعتها.	تميّزت	ميّزت	تميّزوا	فسرت
	Every part of that country is pretty, but the desert was distinguished by its beauty and nature.	was distinguished	differentiated between	were distinguished	explained
24	ليس من الواضح إذا الشاي نشأ في الصين أو في اليابان.	نشأ	أنشأ	نشأوا	ارتدى
	It isn't clear whether tea originated in China or Japan.	originated	founded	originated (pl)	wore
25	مرارا وتكرارا، الأغنية ردّدت البيت الحزين عن الطفل المفقود.	ردّدت	تردّدت	ردّدوا	اكتفيت
	Over and over, the song repeated the sad line about the lost child.	repeated	hesitated	repeated (pl)	was satisfied
26	منذ ألف سنة لم يكن من المعروف إذا الأرض دارت حول الشمس كل سنة أو لا.	دارت	أدارت	داروا	حرصت
	A thousand years ago it was not known whether the Earth revolved around the sun every year or not.	revolved	directed	revolved (pl)	took care
27	على الرغم من أن المصدر الأصلي لم يكن واضحا، القصة بلغت المراسل من مخبر مجهول الاسم في نفس البلدة.	بلغت	أبلغت	بلغوا	ركعت
	Though the original source was unclear, the story reached the reporter from an anonymous informant in the same town.	reached	reported	reached (pl)	knelt
28	على الرغم من التدريب الشامل، الجرح حرّم الرياضي من فرصة المنافسة في السباق المشهور.	حرّم	أحترم	حرّموا	انتخب
	Despite rigorous training, the injury deprived the athlete of the opportunity to compete in the famous race.	deprived	respected	deprived (pl)	elected
29	هذا الصباح في المدرسة، الطفل كرّر كلمات المعلم و قرأ بعض الكتب.	كرّر	تكرّر	كرّروا	اهترأ
	That morning in the school, the child repeated the words of the teacher, and read some books.	repeated	was reiterated	repeated (pl)	frayed
30	كانت القطة تبكي في الشارع، والضجة أيقظت الناس في البيت في الساعة الرابعة.	أيقظت	استيقظت	أيقظوا	اهتمت
	A cat was crying in the street, and the sound woke up the people in the house at four o'clock.	woke (someone)	woke (itself)	woke (pl)	was interested in
31	بعد أن سقطت من على الطاولة، الإبرة اختفت في العشب الطويل.	اختفت	أخفت	اختفوا	اعتزمت
	After it fell from the table, the needle disappeared into the long grass.	disappeared	hid (something)	disappeared (pl)	was determined
32	بداية التسجيل كانت واضحة، ولكن الصوت انخفض وأصبح غير واضح بعد فترة قصيرة.	انخفض	خفض	انخفضوا	غار
	At first the recording was clear, but then the voice got quieter and became indistinct after a short time.	got quieter	lowered (something)	got quieter (pl)	got jealous

33	ليلة أمس في المسرح، الفيلم أمتع الأطفال ولكن آباءهم شعروا بالملل.	أمتع	استمتع	أمتعوا	عرق
	Last night at the theater, the film entertained the children, even though the adults were bored.	entertained	enjoyed	entertained (pl)	sweated
34	خلال الدراسة البحثية، الدواء خفف من درجة حرارة المرضى في المستشفى.	خفف	استخف	خففوا	هرب
	During the research study, the medicine lowered the fevers of the patients in the hospital.	lowered	underestimated	lowered (pl)	fled
35	في شمال البلد، الخريف وصل مبكرا، برغم من أن الصيف كان حارا جدا.	وصل	تواصل	وصلوا	انتبه
	In the north of the country, autumn arrived early, even though the summer before had been very warm.	arrived	pursued	arrived (pl)	paid attention
36	الأسبوع كان مشغول والجدول منع الموظفين من الراحة والاسترخاء.	منع	امتنع	منعوا	أجاد
	The week was busy, and the schedule prevented the employees from resting or relaxing.	prevented	abstained	prevented (pl)	was skilled
37	لم تكن غالية، ولكن الهدية أدهشت المساعدة، ولم تعرف ماذا تقول.	أدهشت	اندهشت	أدهشوا	وشوشت
	It was not expensive, but the gift surprised the assistant, and she didn't know what to say.	surprised	was surprised	surprised (pl)	whispered
38	كان من الواضح أن الوقت مضى بسرعة أكثر خلال عطلة الصيف.	مضى	أمضى	مضوا	انحنى
	It was obvious that time passed more quickly during the summer vacation.	passed	spent (time)	passed (pl)	bent
39	الشهر الماضي في مجلس الشعب، النائب ركّز على سؤال البطالة وعلاجها.	ركّز	ارتكز	ركزوا	انعدم
	Last month in congress, the representative focused on the question of unemployment and how to remedy it.	focused	was arranged in a circle around	focused (pl)	didn't exist
40	لم يصب أي شخص، ولكن الزلزال أخاف الأطفال، الذين ذهبوا تحت مكاتبهم.	أخاف	خاف	أخافوا	حن
	It didn't hurt anyone, but the earthquake scared the children and they got under their desks.	scared	feared	scared (pl)	yearned
41	في اليوم الأخير من الصف، الدرس أعد الطلاب للامتحان النهائي قبل عطلة الشتاء.	أعد	أستعد	أعدوا	قهر
	On the last day of class, the lesson prepared the students for the final exam before winter break.	prepared	got ready	prepared (pl)	won
42	على الطريق إلى المدينة، السيارة جاوزت حد السرعة و السائق دفع غرامة.	جاوزت	أجازت	جاوزوا	نوت
	On the way back to the city, the car exceeded the speed limit and the driver paid a fine.	exceeded	approved	exceeded (pl)	intended
43	العاصفة القادمة صغيرة، ولكن الطقس أقلق البحارة الذين يسافرون في هذا اليوم.	أقلق	قلق	أقلقوا	توسل

	The approaching storm was small, but the weather concerned the sailors who were traveling that day.	concerned	was worried	concerned (pl)	begged
44	التصادم كان مخيف ولكن التجربة علّمت الرجل أن يصبح أكثر حذرا.	علّمت	تعلّمت	علّموا	ارتجفت
	The collision was very frightening, but the experience taught the man to be more careful in the future.	taught	learned	taught (pl)	trembled
45	التفسير كان غريبا، ولكن الدليل أقنع الطبيب أن هذا التفسير كان صحيحا.	أقنع	اقتنع	أقنعوا	انتظر
	The explanation was strange, but the evidence persuaded the doctor that this explanation was correct.	persuaded	was convinced	persuaded (pl)	waited
46	النار انتشرت بسرعة، والمصنع انفجر مع ضجة كبيرة بعد فترة قصيرة.	انفجر	فجّر	انفجروا	عض
	The fire spread quickly and the factory exploded with a loud blast a short while later.	exploded	blew up (something else)	exploded (pl)	bit
47	الرحلة بالسيارة كانت تستغرق عدة أيام، بينما القطار وفّر قدر كبير من الوقت للمسافرين.	وفّر	توفّر	وفّروا	أدعى
	Traveling by car used to take several days, but the train saved travelers a great deal of time.	saved	was fulfilled	saved (pl)	claimed
48	الإصلاحات كانت ضرورية، ومع ذلك التنظيف حسن البيت أكثر من أي عامل آخر.	حسن	استحسن	حسنوا	ذاق
	The repairs were necessary, but the cleaning improved the house more significantly than any other factor.	improved	admired	improved (pl)	tasted
49	لعبة كرة القدم انتهت عندما الكرة ضاعت بين الأشجار ولم يستطع أحد أن يجدها.	ضاعت	أضاعت	ضاعوا	تطوعت
	The soccer game ended when the ball disappeared between the trees and no one could find it.	disappeared	lost (something)	disappeared (pl)	volunteered
50	في نهاية الفصل، الجفاف أمات الزهور التي نبتت على حافة بعيدة من الملعب.	أمات	مات	أماتوا	استعار
	At the end of the season, the drought killed the flowers that were growing at the far edge of the yard.	killed	died	killed (pl)	borrowed
51	القضية كانت خلافية، والجدال استغرق ساعات طويلة، حتى أن المشاركين كانوا متعبين.	استغرق	غرق	استغرقوا	أغمض
	The topic was controversial, so the discussion lasted long hours, until the participants were tired.	lasted	sank	lasted (pl)	blinked
52	كانت هناك كاميرا فوق الباب، وهذه الآلة صورت الزبائن عندما دخلوا الدكان.	صورت	تصورت	صوروا	نضجت
	There was a camera above the door and this machine photographed the customers when they walked into the store.	photographed	imagined	photographed (pl)	ripened
53	بعد سقوطه من على الشجرة ، الورق لمس وجه الرجل النائم، وفاجأته .	لمس	التمس	لمسوا	اعتقل
	As it fell from the tree, the leaf touched the sleeping man's face and surprised him.	touched	asked	touched (pl)	arrested

54	المسار كان مربك، و مع ذلك الخريطة بينت موقع أطلال الحضارة القديمة.	بينت	تبينت	بينوا	استاءت
	The terrain was confusing, but the map indicated the location of the ruins of the ancient civilization.	indicated	appeared	indicated (pl)	resented
55	مخبز جديد فتح على المفترق، والرائحة استهوت كثير من الزبائن في اليوم الأول.	استهوت	هوت	استهووا	استجوبت
	A new bakery opened on the corner and the smell attracted many customers on the first day.	attracted	loved	attracted (pl)	questioned
56	فسر كتاب التاريخ أن التمثال تأسس في هذا المكان منذ عدة سنوات.	تأسس	أسس	تأسسوا	أسف
	The history book explained that the monument was built in this place a number of years ago.	was built	built	was built (pl)	regretted
57	استمر المطر حتى أن الماء كسر السد وغمر جزء من المنطقة.	كسر	انكسر	كسروا	تذمر
	The rain continued to fall until the water broke the dam and flooded part of the area.	broke	broke (itself)	broke (pl)	complained
58	الاقتصاد كان جيّد، والفرصة شجّعت الأجانب أن يأتوا إلى الجزيرة.	شجّعت	تشجعت	شجّعوا	تضايقت
	The economy was good, and the opportunity encouraged foreigners to come to the island	encouraged	was brave	encouraged (pl)	was annoyed
59	كان الجو ممطرا جدا، ولكن الجو أسعد المزارعين في المنطقة.	أسعد	سعد	أسعدوا	رن
	It was very rainy but the weather pleased the farmers in the area.	pleased	was happy	pleased (pl)	rang
60	الفحم كان أحسن، ولكن الحطب عمل أيضا لتسخين الفرن.	عمل	استعمل	عملوا	تسمم
	Coal was better but wood also worked to heat the stove.	worked	used (something)	worked (pl)	poisoned